Prof. univ. dr. Gheorghe-Ioan MIHALAŞ

Doctor în fizică, specialitatea biofizică, la Universitatea Bucureşti, bursă Fulbright la Virginia Commonwealth University, Richmond. Membru titular al Academiei de Ştiinţe Medicale, premiul Gheorghe Marinescu al Academiei Române, preşedinte al Federaţiei Europene de Informatică Medicală (2006-8), titular al cursului de Biostatistică la Studii Doctorale.

Prof. univ. dr. Diana LUNGEANU

Absolventă a Facultăţii de Automatică şi Calculatoare, doctor în ştiinţa calculatoarelor, bursă de specializare în Health Research and Policy la Universitatea Georgetown, activitate în proiecte de prelucrare a datelor medicale.

Într-o lume care respiră tehnică de calcul valoarea unei lucrări didactice de informatică rezidă în diversitatea şi actualitatea temelor abordate fără a renunţa la unitatea stilului şi a formei de prezentare. Lucrarea de faţă s-a cristalizat de-a lungul anilor în spiritul acestor cerinţe, ţinând pasul cu dezvoltarea informaticii medicale. Bogăţia şi actualitatea informaţiilor prezentate este asigurată de calitatea surselor bibliografice: cărţi de specialitate de circulaţie mondială şi resurse Internet de ultimă oră. Relevanţa materialului rezultă dintr-o selecţie judicioasă atemelor abordate, selecţie bazată pe experienţa acumulată de autori prin colaborări internaţionale.

Prof. univ. dr. Adrian NEAGU – referent ştiinţific

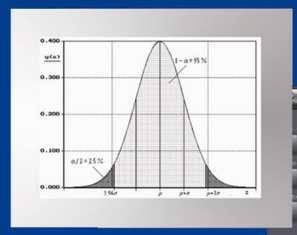
9 789738 775718

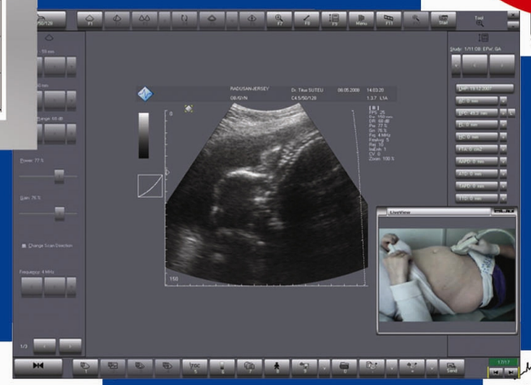
ISBN 978-973-87757-1-8

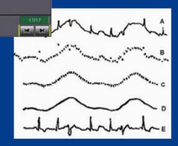
Gh. I. Mihalaș, D. Lungeanu • Informatică Medicală și Biostatistică

Gheorghe Ioan MIHALAŞ
Diana LUNGEANU

Informatică Medicală și Biostatistică







Editura Victor Babeș

Gheorghe-Ioan MIHALAŞ

Diana LUNGEANU

INFORMATICĂ MEDICALĂ ŞI BIOSTATISTICĂ

Colecția

MANUALE

Gheorghe-Ioan MIHALAŞ

Diana LUNGEANU

INFORMATICĂ MEDICALĂ ŞI BIOSTATISTICĂ

2009

Editura VICTOR BABEŞ

Piaţa Eftimie Murgu 2, 300041 Timişoara
Tel./ Fax 0256 495 210
e-mail: evb@umft.ro

Director general: Prof. univ. dr. Ştefan Iosif Drăgulescu
Consilier editorial: Cala Christian

Referent ştiinţific: Prof. univ. dr. Adrian Neagu
Coordonator colecţie: Prof. univ. dr. Andrei Motoc

Editură acreditată de Ministerul Educaţiei şi Cercetării prin
Consiliul Naţional al Cercetării Ştiinţifice din Învăţământul Superior, cod 324.

Descrierea CIP a Bibliotecii Naţionale a României
MIHALAŞ, GHEORGHE-IOAN
Informatică medical şi biostatistică / Gheorghe- Ioan Mihalaş, Diana Lungeanu. - Timişoara: Editura Victor Babeş, 2009
Bibliogr.
ISBN 978-973-87757-1-8

I. Lungeanu, Diana

004:61(075.8)
519.22:57(075.8)

Tipărit la Tipografia Eurostampa
B-dul Revoluţia din 1989 nr. 26, Timişoara
Tel. 0256- 204 816, edituraeurostampa@gmail.com

CUPRINS

Partea a III-a

SEMNALE ŞI IMAGINI BIO-MEDICALE

Partea a IV-a

DECIZIA MEDICALA ASISTATA DE CALCULATOR

Partea I

NOȚIUNI GENERALE

OBIECTUL INFORMATICII MEDICALE

Informatica medicală este o disciplină tânără, termenul apărând în cursul anilor ’60. În accepțiunea inițială informatica medicală cuprindea *programele de calculator* cu aplicabilitate în domeniul medical. Progresul tehnic rapid a arătat însă că, pentru aceleaşi aplicații, atât programele cât şi suportul fizic se schimbau; ceea ce rămânea la fel era modul în care era prelucrată *informația.*

Astfel, în accepțiunea actuală, centrul definiției s-a mutat de la „calculator” la informație. Coiera [1997] chiar atrage atenția în acest sens: „Informatica medicală se ocupă de calculatoare tot atât de mult cât se ocupă cardiologia de stetoscoape”.

Definiții: Obiectul informaticii medicale – caseta 1.1.

TEORIA INFORMAȚIEI

Noțiunea de informație

Pentru a ne ocupa de *informația medicală*, să încercăm mai întâi să privim conceptul de *informație* la modul general.

Termenul de informație este folosit în mod curent în viața de zi cu zi, fiind cel mai adesea asociat cu aducerea unui element de noutate. Fiind un concept cu grad înalt de generalitate (categorie filosofică), informația nu poate fi definită în manieră clasică, pornind de la genul proxim şi precizând diferențele specifice, ci prin propietatea sa esențială – cea de a înlătura o nedeterminare.

Noțiunea de informație – caseta 1.2.

Proprietățile informației

Informația nu este materie; totuşi ea nu poate exista înafara materiei. Norbert Wiener spunea “Creierul nu secretă informație precum ficatul fiere”.

Informația nu este energie; totuşi ea nu se poate transmite fără un suport energetic.

Nu este o relație directă între cantitatea de energie ce însoțește transmiterea unei informații şi cantitatea de informație transmisă. De exemplu, energia unui trăsnet în timpul unei furtuni este imensă, însă informația transmisă este neglijabilă; în schimb un foşnet într-o pădure, purtat de o energie infimă poate reprezenta o informație vitală – punând pe fugă un animal! Să semnalăm totuşi că nu există relație nici între cantitatea de informație şi efectele sale. De ex. Legenda lui Tezeu din metodologia greacă: Tezeu promisese tatălui său Aegeus, că dacă va invinge în luptă minotaurul va înlocui pânza neagră a corabiei cu pânză albă, dar a uitat şi tatăl său s-a aruncat de pe stânci. Informația primită a fost doar 1 bit.

Triada abordărilor complete

Introducerea aspectelor informaționale în studiul materiei vii completează imaginea noastră privind complexitatea sistemelor biologice, actualmente considerându-se că o abordare completă trebuie să acopere atât aspectele materiale şi energetice cât şi cele informaționale.

Triada abordărilor complete – caseta 1.3.

Valoarea utilă a informației

Sensul noțiunii de informație, așa cum a fost prezentat mai sus etc. legat de altă noțiune – nederminarea (sau incertitudinea) – vag definită la rândul său. Acelaşi mesaj poate să aibă valori informaționale diferite pentru diferiți receptori: pentru o persoană care deja ştia conținutul său cantitatea de informație primită este zero, însă pentru receptorii care nu-i ştiau conținutul va putea fi evaluată cantitatea de informație primită deci:

Valoarea utilă a informației depinde de receptor.

Caseta 1.1

Obiectul informaticii medicale

Accepţiunea clasică: totalitatea programelor de ***calculator*** cu aplicaţii în domeniul biomedical şi sănătate.

Definiţia actuală: disciplina care studiază întregul flux al ***informaţiei medicale:*** generare, achiziţie, stocare, transmitere, prelucrare şi utilizare.

Caseta 1.1a

Structura cursului de informatică medicală

Partea I **-** Nivel individual

Secţiunea A. Date

1. Stocare: Baze de date medicale
2. Prelucrare
   - *2a: Date calitative şi numerice:* Biostatistica
   - *2b: Date grafice:* Semnale biologice
   - *2c: Imagini*: Imagistica medicală

Secţiunea B. Cunoştinţe

3. Decizia medicală asistată de calculator

Partea II - Nivel supraindividual

4. Sisteme informatice în sănătate

Caseta 1.2

Noţiunea de informaţie

Informaţia este un concept cu grad înalt de generalitate caracterizat prin proprietatea de a înlătura o nedeterminare (incertitudine).

Caseta 1.3

Triada abordărilor complete

- aspectul material – structura
- aspectul energetic – suportul funcţional
- aspectul informaţional – mecanismul funcţional

Valoarea utilă a informaţiei depinde de receptor

Caseta 1.4

Cantitatea de informaţie

- Cantitatea de informaţie eliberată de un eveniment a carui probabilitate este p_i

$$I_i = -\log_2 \mathrm{p_i} \tag{1}$$

- Unitatea de măsură pentru cantitatea de informaţie = bit
- *Definiţie:* Un bit este cantitatea de informaţie primită când se înlătură o nedeterminare de 1/2
- Pentru o succesiune de N evenimente (mesaj de lungime N)

$$I = \sum_{i=1}^{k} n_i I_i \tag{2}$$

- Entropia informaţională este cantitatea medie de informaţie per eveniment (simbol) într-un mesaj:

$$H = -\sum_{i=1}^{k} p_i \log_2 p_i \tag{3}$$

Caseta 1.5

Redondanţa

- Entropia maximă: pentru evenimente echiprobabile $p_i = 1 / k$, de unde

$$H_{max} = log_2 k \tag{1}$$

- Redondanţa absolută:

$$R = H_{max} - H_{real} \tag{2}$$

- Redondanţă relativă:

$$R_r = R / H_{max} \tag{3}$$

Cantitatea de informație

Parcurgerea acestui subiect necesită cunoştinţe fundamentale de teoria probabilităților

Pornind de la proprietatea fundamentală a informației, cea de a înlătura o nedeterminare, Shannon a considerat că informația primită este invers proporțională cu probabilitatea de apariție a evenimentului: dacă se va întâmpla un eveniment cu probabilitate mare, informația primită este mică; în schimb primim o informație "mai mare" dacă apare un eveniment mai rar. (Ziariştii exploatează intens această relație!)

Relația propusă de Shannon pentru calculul cantității de informație care este primită când se petrece un eveniment cu probabilitatea p_i cuprinde logaritmul în baza 2 din inversul probabilității p_i (formula (1) în caseta 1.4).

Pe baza acestei relații stabileste unitatea de măsură pentru cantitatea de informație, numită *bit* (de la BInary digiT).

În mod usual informația se transmite printr-o succesiune de evenimente, numită adesea *mesaj*, iar un eveniment într-un mesaj se mai numeşte *simbol*.

În cazul unui mesaj format din N evenimente, fiecare eveniment *i* apare de n_i ori, aducând de fiecare dată informația I_i, deci mesajul aduce informația

$$I = n_1 I_1 + \eta_2 I_2 + ... + n_k I_k = \sum_{i=1}^{k} n_i I_i$$ relație care este trecută şi în caseta 1.2.

Valoarea medie a informației corespunzătoare unui eveniment într-un şir de N evenimente se mai numeşte "entropie informațională, H, şi se calculează astfel:

$$H = \frac{n_1 I_1 + n_2 I_2 + ... + n_k I_k}{N} = \frac{n_1}{N} I_1 + \frac{n_2}{N} I_2 + ... + \frac{n_k}{N} I_k$$

la limită (i.e. atunci cand $N \to \infty$) relația devine:

$$H = p_1 I_1 + ... + p_k I_k = \sum_{i=1}^{k} p_i I_i$$

Înlocuind I_i conform relației (1), obținem formula (3) din caseta 1.4, formulă fundamentală în teroria informației, numită şi formula lui Shannon pentru entropia informațională.

Relația între entropia informațională şi entropia termodinamică

Termenul de "entropie" a fost introdus în termodinamică pentru enunțarea principiului al II-lea al termodinamicii: "În procesele termodinamice entropia nu poate să scadă: ea rămâne constantă în cadrul proceselor reversibile şi creşte în cazul proceselor ireversibile".

Relația între entropia termodinamică şi cea informațională poate fi înțeleasă pornind de la experimentul "ideal" propus de Maxwell pentru explicarea variației entropiei în cazul proceselor ireversibile, prezentat în cadrul cursului de biofizică.

Se vede deci că sistemul poate evolua în sens contrar celui dictat de al II-lea principiu al termodinamicii în cazul în care primeşte o informație. Acesta este mecanismul prin care sistemele vii evoluează spre stări tot mai organizate şi deosebite de mediul înconjurător.

Redondanță

Entropia informațională are valoare maximă când evenimentele din mesaj sunt echiprobabile: $p_i = \frac{1}{k}$;

atunci $H_{\max} = k\left(\frac{1}{k}\log\frac{1}{k}\right)$ de unde se obține relația (1) din caseta 1.5.

Un exemplu ar fi cazul unui mesaj encriptat, în care probabilitatea apariției unui simbol este (cel puțin aparent) independentă de simbolurile anterioare. În mesajele reale probabilitatea unui simbol depinde de simbolurile anterioare; putem, în funcție de context, să "ghicim" ce urmează, putem folosi prescurtări, putem observa greşeli cum ar fi omisiunea unei litere etc. Deci informația nu este distribuită uniform în mesaj sau chiar în interiorul cuvintelor, cantitatea de informație transportată în realitate fiind inferioară celei maxime ce ar putea fi transmise folosind aceeaşi lungime a textului. Această diferență, între cantitatea maximă ce poate fi conținută în mesaj şi cea reală se numeşte redondanță şi reprezintă o parte din mesaj care… nu conține informație!

Relația de definiție a redondanței absolute este:

$R = H_{max} - H_{real}$

Raportând redondanța absolută la H_{max} se defineşte redondanța relativă (vezi caseta 1.5).

Rolul redondanței

Aparent redondanța ar reprezenta o încărcătură inutilă în mesaj. Totuşi, prezența ei diminuează rolul negativ al perturbațiilor ce apar în cursul transmiterii informației, putând deseori reconstitui mesajul inițial chiar dacă unele simboluri au fost perturbate.

Transmiterea informației

Am precizat anterior că valoarea utilă a informației depinde de receptor, deci noțiunea de informație are sens doar dacă se transmite; altfel, rămâne în faza de "informație potențială".

Transmiterea informației presupune o *sursă* a informației (emițător E) şi un destinatar (receptor R). Spațiul dintre S şi R reprezintă *canalul de comunicație* (C). Pe canalul de comunicație pot să apară diverse *zgomote* care perturbă sistemul de comunicație afectând calitatea transmisiei.

Să introducem doi termeni importanți în cadrul sistemelor de comunicație:

- *mesaj* – un termen pe care îl folosim când ne referim la conținutul informațional al transmisiei
- *semnal* – suportul fizic care transportă mesajul (sunet, current electric etc.).

Pentru diminuarea efectelor perturbațiilor sau pentru a asigura transmiterea mesajului la distanțe foarte mari, se introduc pe canalul de transmisie nişte dispozitive numite *traductori*. Un traductor schimbă suportul fizic al unui semnal. De exemplu, în cazul unei convorbiri telefonice, microfonul este traductorul localizat lângă emițător, transformând sunetele (variații ale presiunii aerului) în variații ale unui curent electric. Canalul de comunicație este reprezentat de firele telefonice. La destinatar un alt traductor, casca telefonică, transformă variațiile curentului electric în vibrații ale unei membrane elastice generând astfel sunete.

Există şi alte dispozitive ce pot fi utilizate în sisteme de comunicație, de ex. *modem-ul*. Denumirea modem provine de la mod<u>ulator/dem</u>olator. Modemul este un dispozitiv care asigură modularea semnalului, adică suprapunerea semnalului real peste un semnal purtător (undă purtătoare) care are caracteristici încât se diminuează efectul perturbațiilor (de ex. perturbațiile uzuale, de joasă frecvență, sunt eliminate dacă unda purtătoare are frecvență înaltă).

O altă transformare pe care o putem aplica semnalului pentru transmisie este *codificarea*. Mesajul este compus uzual dintr-o succesiune de simboluri. Totalitatea simbolurilor utilizate pentru a compune un mesaj se numeşte *alfabet*. Simbolurile

alfabetului se mai numesc "*litere*", iar cu literele putem construi *cuvinte.* Totalitatea cuvintelor cu sens reprezintă un *dicționar*, iar precizarea sensului cuvintelor se numeşte *semantică*. Cu ajutorul cuvintelor se pot construi propoziții; regulile de construcție a propozițiilor se numeşte *sintaxă.* Un dicționar împreună cu semantica şi o sinteză reprezintă un *limbaj*. Noi folosim uzual pentru comunicație *limbaje naturale*, dar există posibilitatea utilizării unor *limbaje formale* sau *artificiale.* Diferitele componente ale sistemului de comunicație pot folosi diferite alfabete sau dicționare. Transpunerea unui mesaj dintr-o formă ce utilizează un alfabet într-o formă în alt alfabet, cu anumite reguli de corespondentă se numeşte *codificare*. Operațiunea inversă se numeşte *decodificare.* Transpunerea unui mesaj dintr-un limbaj în altul se numeşte *traducere*.

Să mai menționăm legat de sistemele de comunicație că există o capacitate limitată de transmisie a informației pe canalul de comunicație, numită *viteză de transmisie*, măsurată în bit/secundă.

Exemple de transmisie a informației în materie vie

a) *Codul genetic*. Informația privind structura proteinelor ce pot fi sintetizate este stocată în molecula de AND din nucleu. Acizii nucleici conțin 4 baze azotate: adenina A, timina T, citozina C şi guanina G (în cazul ARN în loc de timină apare uracilul u). Proteinele sunt formate din 20 de aminoacizi esențiali. O succesiune de 3 baze azotate din AND se numeşte *codon* şi poartă informația pentru codificarea unui aminoacid într-o secvență proteică. Totalitatea corespondenților între codoni şi aminoacizii corespunzători poartă denumirea de *cod genetic*. Porțiunea dintr-un lanț AND care poartă informația pentru sinteza unei proteine se numeşte *genă*, iar ansamblul tuturor genelor unei specii se numeşte genom. Genomul uman conține circa 30.000 gene.

Exercițiu: Ce cantitate medie de informație poartă un aminoacid într-o structură proteică având 100 aminoacizi?

Rezolvare: *Considerăm că cei 20 aminoacizi au aceeaşi probabilitate de apariție într-o secvență proteică (ipoteză relativ depărtată de realitate, dar simplificatoare pentru rezolvarea problemei). Calculăm entropia informațională cu relația (3) din caseta 1.4 înlocuind $p_i = 1/20$, deci:*

$$H = -\sum_{1}^{20} (1/20)\log_2 (1/20) = \frac{100}{20}\cdot \log_2 20 \approx 5\cdot 4{,}2 = 21\,biti$$

b) *Codificarea informației în sistemul nervos.* Pe axoni informația este transmisă printr-o succesiune de impulsuri nervoase; fiecare impuls nervos este un potențial de acțiune care are întotdeauna aceeaşi amplitudine. Unui stimul mai intens îi corespunde o rată mai ridicată de formare a potențialelor de acțiune; spunem că pe axon informația privind intensitatea stimulului este *codificată în frecvență*.

La nivelul sinapselor are loc o descărcare a veziculelor cu mediator chimic în spațiul sinaptic, cantitatea de mediator descărcată fiind proporțională cu frecvența impulsurilor nervoase pe axon; spunem că în spațiul sinaptic informația privind intensitatea stimulului este *codificată în amplitudine*, aceasta fiind reprezentată de cantitatea de mediator descărcată.

La nivelul membranei postsinaptice, mediatorul se cuplează pe receptorii postsinaptici, se deschid canalele de sodium, membrana se depolarizează şi apare un potențial care se propagă pe membrana corpului neuronal sau pe dendrite. Spunem că informația este *codificată în amplitudine*, aceasta fiind reprezentată de potențialul local.

Informatica medicală

După această incursiune în teoria informației putem reveni la noțiunea centrală din informatică medicală şi anume *informația medicală*.

Ce este informația medicală şi când apare ea?

Date şi cunoştințe

Să încercăm să schițăm în cel mai simplificat mod actul medical primar şi anume vizita pacientului la medic. Poziția centrală în activitate medicală este ocupată de *pacient*. Fără pacient nu există medicină! Actorul principal al activității medicale este *medicul,* dar în activitatea medicală sunt implicate numeroase alte persoane care aparțin aşa-numitelor "*profesii aliate*". Dialogul medic-pacient începe uzual cu expunerea de către pacient a motivelor pentru care s-a prezentat la medic. Această descriere reprezintă transmiterea unor informații de la pacient către medic. Informațiile care se transmit sau se utilizează într-un act medical (sau ca urmare a unui act medical) reprezintă *informația medicală*. Dialogul este succedat de către examenul obiectiv al pacientului, medicul colectând astfel şi alte informații despre pacient. Să observăm că aceste informații au un caracter individual – sunt valabile pentru *acest pacient.* Aceste informații se numesc ***date***. Uzual paleta datelor se completează cu informații provenind şi din alte investigații (probe de laborator, explorări funcționale, radiografii etc.). Indiferent cât de complexe ar fi ca reprezentare, ele sunt "date", fiind caracteristice unui anumit individ.

În acelaşi timp, medicul utilizează şi alt fel de informații, numite *cunoştințe*. Acestea au un caracter general şi sunt acumulate în cursul pregătirii profesionale precum şi în experiența sa practică. Fără aceste cunoştințe informațiile sub formă de date nu pot fi interpretate (revenim la afirmația că valoarea utilă a informației depinde de receptor; practic, fără aceste cunoştințe receptorul datelor nu este "medic"). De aceea numeroşi autori numesc *informație* doar *datele interpretate*. Pentru a evita confuzia între termenul *informație* folosit la modul general şi *informație* pentru treapta de *date interpretate*, vom păstra termenul de *date interpretate* pentru acest nivel.

Ciclul elementar al informației medicale

Prin interpretarea datelor de către medic pe baza cunoştințelor sale, este generată o nouă informație numită diagnostic. Pe baza diagnosticului, folosind din nou cunoştințele sale, medicul stabileşte un plan terapeutic pe care îl aplică pacientului cu scopul de a îmbunătății starea pacientului. Urmărirea evoluției pacientului este însoțită de colectarea unor noi informații sub formă de date. Se observă că se închide un ciclu al fluxului informațional în activitatea medicală, numit "ciclul elementar al informației medicale".

Tipuri de date

Informațiile culese despre starea pacientului, adică datele, pot îmbrăca diverse forme:

- *date calitative* – cu caracter descriptiv, aşa cum apar în anamneză
- *date numerice* – forma uzuală de prezentare a rezultatelor de laborator
- *grafice* – modul de înregistrare a evoluției în timp a unor mărimi biologice (ex.: semnalul ECG, EEG etc.)
- *sunete* – de ex. fonocardiograma; modul de prelucrare este asemănător cu cel al altor semnale
- *imagini* – radiografia, tomografia, ecografia etc.
- *imagini dinamice* – filme.

Modul de achiziție, stocare şi prelucrare este specific pentru fiecare tip de date şi în cadrul cursului nostru le corespund capitole separate.

Tipuri de cunoştințe

Cunoştințele pot fi de mai multe feluri:

- *cunoştințe explicite* – care se pot formaliza, se pot exprima în propoziții, pot fi uşor transmise pe cale orală sau scrisă
- *abilitati* sau *cunoştințe tacite* (limba engleza - *skill*) – cele câştigate prin experiență practică (de ex. îndemânarea unui chirurg sau a unui dentist); nu pot fi transmise uşor.

Clasificarea informației medicale pe nivele structurale

În ciclul elementar al informației medicale prezentat mai sus am luat în considerare informațiile care apar în activitatea medicală curentă, la nivelul individului, numit pacient. Totuşi fenomenele care se petrec în materia vie (legate de starea de sănătate a pacientului) privesc deseori nivele infraindividuale, pornind de la nivelul molecular sau celular, urcând prin nivelul de țesut, organ sau sistem până la nivelul întregului organism sau nivelul individual.

Pe de altă parte, activitatea medicală este organizată în unități care prestează servicii pentru populație, deci putem urmări fluxul informațional şi la nivel supraindividual, de comunitate. Corespunzător acestor nivele structurale avem diferite discipline biomedicale precum şi diferite capitole corespunzătoare ale informaticii medicale.

Operații cu informații

Urmărind ciclul de viață al informației, din momentul generării sale până în momentul utilizării, observăm că informația suferă o serie de operații:

- *achiziția (colectarea)* – presupune mijloace specifice tipului de informație
- *stocarea* – baze de date, respective baze de cunoştințe
- *transmitere* – căi, procedee
- *prelucrare* – cu o largă paletă de metode specifice, pentru a extrage elementele esențiale în vederea interpretării şi utilizării
- *protecție* – măsurile ce se impun pentru asigurarea integrității informației stocate sau transmise, precum şi a confidențialității acesteia
- *interpretare/utilizare* – pasul final, în care informația este integrată în acțiunile specifice nivelului.

Capitolele informaticii medicale şi structura cursului

Structura schematică a cursului cu durata de un semestru (predat în anul I la studenții facultății de medicină) este prezentată în caseta 1.1a. Partea referitoare la cunoştințele medicale este prezentată sumar, la nivel introductiv.

Noțiunile de modelare, bioinformatică şi neuroinformatică se predau numai sub formă de cursuri avansate şi nu sunt cuprinse în acest volum.

BAZE DE DATE MEDICALE. NOȚIUNI INTRODUCTIVE

Ce sunt bazele de date ?

Sunteţi mult mai familiarizaţi cu acest concept decât credeţi. Întâlniţi baze de date în viaţa de fiecare zi. Aţi răsfoit vreodată un program TV? Aţi consultat un dicţionar sau o enciclopedie? Aţi intrat într-o bibliotecă? Ei, toate astea sunt baze de date.

O baza de date este o colecție organizată de date. O bază de date de calculator va fi o colecție de date organizată în calculator (mai exact, în fișiere). Ce fel de date? Orice fel: liste cu nume şi adrese, cărțile dintr-o bibliotecă, orice doriți să organizați şi să păstrați.

1. FIŞIERE DE DATE

DEFINIŢII

Fişier (*file*) = o colecţie organizată de date

Date *(data*) = reprezentări formalizate sau fapte (“instante”), adecvate prelucrarilor umane sau automate

Ce nevoie am de “fişiere de date”? Nu pot înregistra orice informaţii într-un fişier creat cu un editor de texte, de exmplu cu Word?!

NU! Categoric nu.

Un editor de text crează fişiere de tip text şi ne oferă facilităţi de redactare şi aranjare a textului destinat afişării şi/sau tipăririi.

În fişierele de date informaţia este organizată după o *schemă*: anumite structuri sau machete precizate de noi, astfel încat să regăsim uşor orice informaţie odată înscrisă în baza de date şi să o putem prelucra după dorinţă.

Un fişier de date în format electronic (“pe calculator”) nu va fi o structură rigidă ca o carte de telefon sau un mers al trenurilor tipărite pe hârtie. El ne oferă o structură flexibilă prin faptul că putem căuta datele după diverse criterii şi le putem chiar modifica ordinea. Putem lua ca exemplu fişierul de date reprezentat de catalogul cu fişe dintr-o bibliotecă, care are un anumit format fix. Deoarece fiecare carte este afişată prin titlul ei sau prin autor, pentru a găsi o anumită carte din acel catalog, trebuie să

cunoaştem titlul sau autorul. Dacă nu ne putem aminti cu suficienta precizie nici una din aceste informaţii, rezultatul va fi o căutare întortocheată prin toate fişele catalogului. Dacă acest catalog ar fi fost un fişier de date "pe calculator", am fi putut căuta nu numai după nume sau autor, ci şi după diverse cuvinte-cheie, după data publicării, sau după frânturile de informaţie despre autor pe care ni le amintim.

Un lucru foarte important este SCHEMA sau STRUCTURA UNUI FISIER DE DATE:

Element de structură	**Nume în lb. engleză**	**Explicaţie, exemplu**
înregistare	*record*	similarul unei fişe clasice (fişa de carte la bibliotecă)
câmp	*field*	similarul unei rubrici din fişa clasică ("numele şi prenumele", "sex", "data naşterii", etc.)
articol	*item*	conţinutul concret al unui câmp, respectiv al unei rubrici ("POPESCU ALEXANDRU" atunci când reprezintă conţinutul câmpului *numele şi prenumele*)

Dacă ne propunem să ţinem evidenţa pacienţilor folosind un calculator, vom înscrie datele care erau conţinute în fişa de evidenţă clasică (pe hârtie) în înregistrări conţinute în fişiere de date, aşa cum ne propun figurile I.1 si I.2.

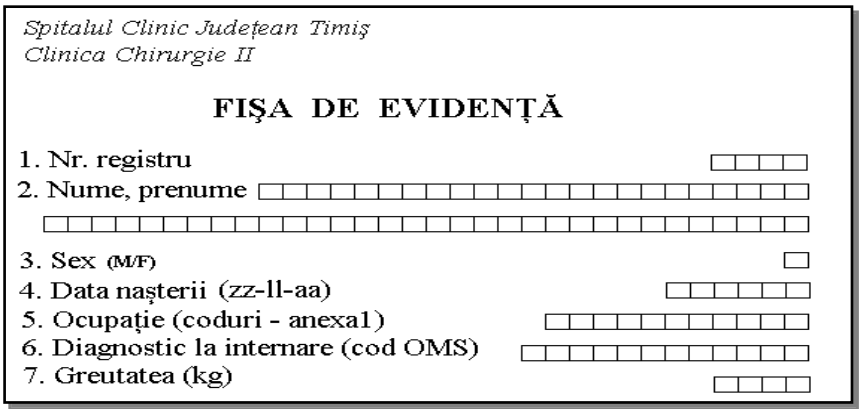


Figura I.1. Exemplu simplificat de fişa clasică

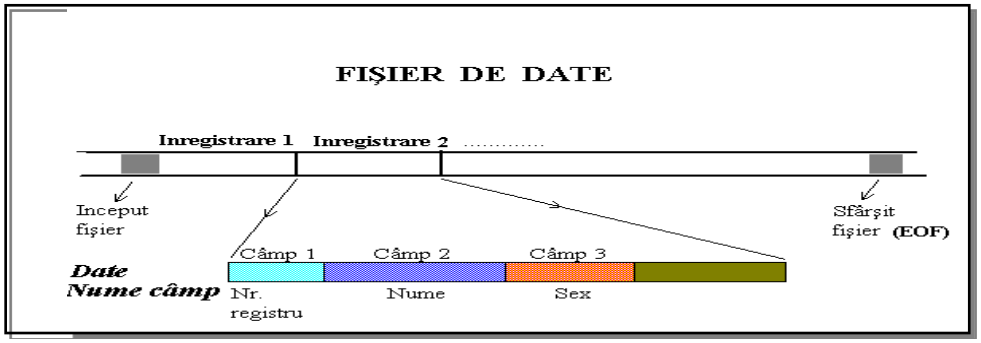


Figura I.2. Organizarea unui fişier de date secvenţial

Ne putem imagina fişierul de date ca fiind un dosar cu înregistrări (fişe de evidenţă, în accepţiunea clasică). Observaţi că fiecărui pacient îi va corespunde o înregistrare (o fişă de evidenţă) - *record*. Fiecare înregistrare este formată din câmpuri (rubricile ce trebuiau completate la o fişă clasică) - *fields*. Fiecare câmp are un nume şi o dimensiune (există rubrici mai "încapătoare", ca cele pentru *nume*, *diagnostic* şi unele mai "înguste", ca cele pentru *sex* sau *greutate*).

Datele concrete care se introduc la un moment dat într-un anume câmp le vom numi articole - *items* (ca de exemplu, POPESCU pentru *nume*, sau M pentru *sex*, etc.).

Fiecare câmp se va caracteriza prin câteva proprietăţi:

Proprietate câmp	**Exemplu**
nume câmp	"Nume", "Data naşterii", "Sex", etc.
tip câmp	numeric (intreg, real), caracter, logic, data calendaristică, etc.
dimensiune câmp	50 caractere, numar real cu 3 cifre la partea întreagă şi 2 cifre la cea zecimală, etc.

Toate aceste elemente de structură trebuiesc definite la crearea fişierului de date.

Putem deci vedea fişierele de date ca nişte tabele pe care le definim atunci când precizăm structura şi le "umplem" apoi cu datele propriu-zise (cu articole sau *items* concrete) - figura I.3.

Nr. reg.	Nume	Sex	Data_naşt.	Ocupaţie	Diagnostic	Greutatea
2345	Ionescu Adrian	M		.....	.....	.....

Figura I.3. Organizarea logică a unui fişier de date sub forma unui table

2. BAZE DE DATE

O bază de date este formată din unul sau mai multe fişiere de date, dar este mai mult decât o simpla colecţie de fişiere: include, pe lângă acestea, descrierea relaţiilor dintre înregistrări, descriere apelată şi utilizată pe toată durata prelucrării informaţiilor. Figura I.4 prezintă comparativ organizarea informaţiei sub forma unei baze de date la nivel instituţional versus o colecţie de fişiere independente, corespunzătoare diferitelor departamente dintr-o instituţie.

DEFINIŢIE

Baza de date *(database)* = o înregistrare structurată de date: conţine atât datele, cât şi *schema* acestora, adică mijloacele de a stabili şi a menţine relaţii între date; aceste relaţii trebuie să reflecte relaţiile dintre entităţile reale descrise de date (obiecte fizice, evenimente, concepte abstracte)

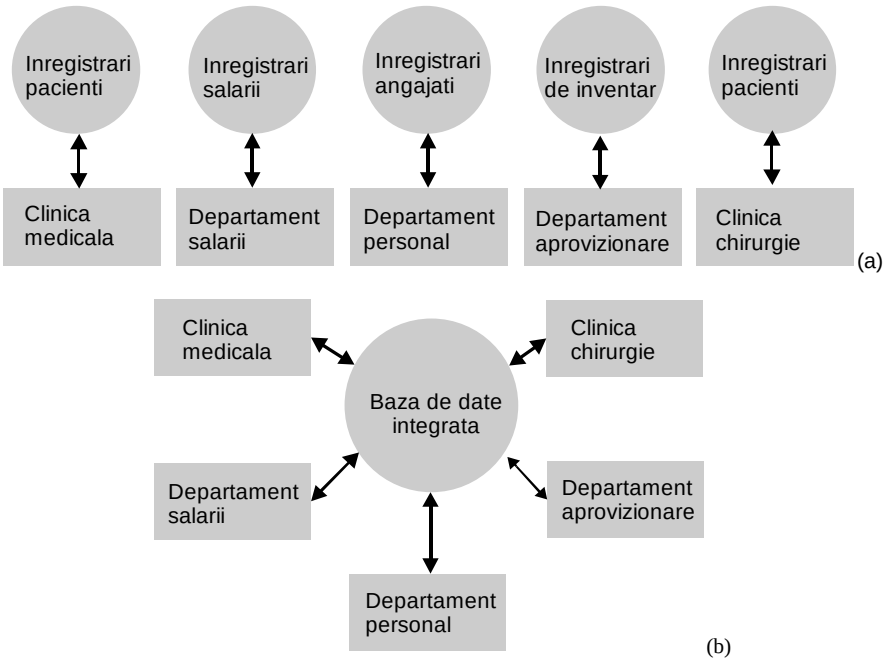


Figura I.4. Organizarea informației într-o bază de date la nivel instituțional (b) comparat cu existența unor fișiere de date independente (a) [Brookshear 2007]

La culegerea datelor trebuie neaparat să fie deja stabilită structura fişierelor şi criteriile de validare, care trebuie să țină seama de alcătuirea fişei de evidență clasice (cu cât discrepanțele de organizare sunt mai substanțiale, cu atât mai mari vor fi şansele de eroare la introducerea datelor şi reținerea personalului în adoptarea evidenței electronice a informatiei). Pentru compactarea datelor din fişiere se pot folosi codificari.

Am auzit de "validarea datelor". Ce înseamnă asta?

Validarea datelor la introducerea lor în baza de date o putem vedea pe mai multe nivele. Există o validare primară, care se face implicit prin modul în care a fost precizată structura: nu voi putea introduce caractere alfabetice într-un câmp numeric, aşa cum, dacă voi introduce cifre într-un câmp de tip caracter, ele nu vor avea nici un fel de semnificaţie valorică (nu voi putea face operaţii matematice cu ele). Pe al doilea nivel de validare există posibilitatea precizării unor anumite valori sau intervale de valori pe care le pot lua articolele din câmpuri: de exemplu, "M" şi "F" pentru câmpul *sex*, valori pozitive şi mai mici decât 2.5 pentru câmpul *înălţime*, etc. Poate exista şi un al treilea nivel de validare, în care se iau în considerare criterii mai complexe, care să ţină cont de eventualele relaţii între câmpuri şi înregistrări.

Dezvoltarea schemei unei baze de date se face în etapa de proiectare sau de *design* a acesteia. Odată definită schema sau structura bazei de date, utilizatorul nu va mai fi preocupat de chestiuni legate de organizarea fizică a datelor în fişierele care

compun baza de date. De asta se vor ocupa programe speciale. Utilizatorul va face referire la date prin numele câmpurilor şi astfel programele scrise pentru consultarea şi administrarea bazelor de date vor fi independente de configuraşia fizică.

3. SISTEME DE GESTIUNE A BAZELOR DE DATE

DEFINIŢIE

Sistem de Gestiune a Bazelor de Date - SGBD (*DBMS - DataBase Management System)* = un set de "unelte *software*" corelate ce au ca scop "construcţia" unei baze de date şi apoi accesul la aceasta; în plus, ele controlează securitatea, integritatea şi secretul datelor

Aceste instrumente *software* încorporează suplimentar: limbaje specializate pentru descrierea şi manipularea datelor; (eventual) un sistem de dicționare de date. Accesul şi manipularea datelor se poate face direct prin funcțiile oferite de SGBD. Cel mai adesea însă, acest lucru îl fac utilizatorii specializați. Utilizatorul obişnuit va utiliza baza de date la nivelul aplicațiilor specifice scopului dorit, prin intermediul unor interfețe speciale definite pentru conectarea programelor de aplicație la DBMS (*API – Application Programming Interface*). Figura I.5 prezintă acest concept de acces pe mai multe niveluri la baza de date propriu-zisă.

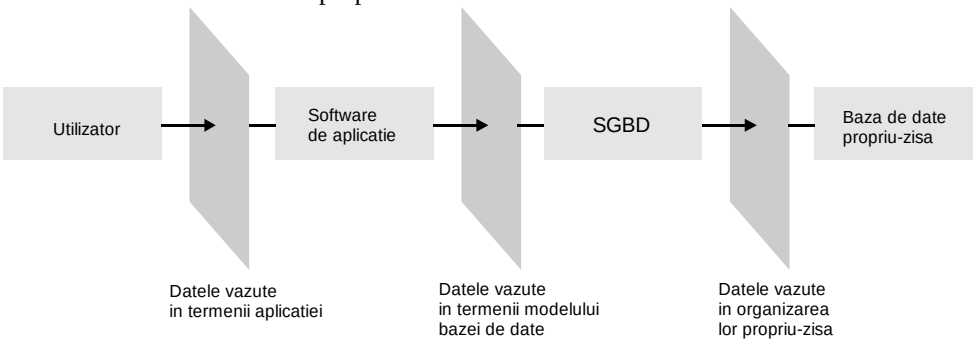


Figura I.5. Conceptul de stratificare a accesului la baza de date: poziția utilizatorului față de baza de date [Brookshear 2007]

Un SGBD are trei funcții de bază:

- funcția de descriere
- funcția de manipulare
- funcția de utilizare.

FUNCȚIA DE DESCRIERE

Funcția de descriere permite descrierea schemei bazei de date (structura datelor şi relațiile dintre acestea). Totodată, se definesc şi condițiile de acces la informațiile conținute în baza de date. Descrierea schemei se realizează cu ajutorul unui limbaj de descriere propriu fiecărui sistem de gestiune.

FUNCȚIA DE MANIPULARE

Funcția de manipulare permite efectuarea următoarelor operații: crearea, inserarea, suprimarea sau actualizarea unor înregistrari definite de utilizator; facilitează căutarea, sortarea şi editarea totală sau parțială a unor înregistrări corespunzătoare rezultatului unei întrebări formulate în acest limbaj.

Limbajele de manipulare pot fi grupate în două mari categorii:

- **limbaje autonome** - sunt de sine stătătoare, în cadrul lor comenzile de manipulare reprezintă chiar funcții referitoare la utilizarea datelor;
- **limbaje care au nevoie de limbaj gazdă** - oferă doar facilitățile de structurare şi căutare, celelalte operații legate de manevrarea fişierelor şi de prelucrare a datelor fiind realizate cu ajutorul unor limbaje de nivel înalt universale.

FUNCȚIA DE UTILIZARE

Funcția de utilizare permite comunicarea între utilizator şi baza de date (sub aspectul asigurării acelor mijloace şi servicii care îl avantajează cel mai mult pe utilizator).
Din punct de vedere al funcției de utilizare, utilizatorii pot fi de mai multe categorii:

- **utilizatori liberi** sau **conversaționali**, care au la dispozitie limbaje de interogare într-o formă foarte apropiată de vorbirea curentă şi formează grupa utilizatorilor aşa-zişi nespecialişti. Întrebările sunt prestabilite, utilizatorii nu cunosc structura sau modul de lucru cu baza de date şi se rezumă la apelarea unor proceduri sau programe corespunzătoare anumitor aplicații;
- **utilizatorii parametrici** fac uz, de regulă, de limbajele de manipulare (în special pentru interogare) utilizând proceduri prestabilite. Ei au cunoştințe de programare şi cunosc atât structura bazei de date cât şi problemele sistemului de operare;
- **administratorul bazei de date** este un utilizator special, care răspunde de toate activitățile şi operațiile referitoare la baza de date pe care o gestionează, inclusiv performanțele acesteia. El defineşte obiectivele sistemului, ajută la definirea cerințelor utilizatorilor, defineşte structura virtuală şi împarte drepturile de acces ale utilizatorilor, stabileşte procedurile de validare a datelor, elaborează concepția de protecție a datelor şi evaluează performanțele sistemului. Administratorul răspunde de alegerea şi implementarea SGBD-ului, asigură încărcarea bazei de date, defineşte strategia de lucru şi distribuie documentația tuturor utilizatorilor.

Pentru orice aplicație de baze de date de dimensiuni medii sau largi există cel putin o persoană cu sarcini de administrare. Aceste persoane (i.e. *administratorii bazei de date*) stabilesc politicile de utilizare şi au grijă ca ele sa fie respectate. Funcțiile specifice de administrare pe care o SGBD trebuie sa le ofere sunt: (i) start/stop aplicația de baze de date; (ii) funcții pentru definirea grupurilor de utilizatori şi pentru controlul accesului; (iii) funcții de arhivare, salvare şi restaurare; (iv) funcții de control al securității şi integrității datelor; (v) importul/exportul datelor; (vi) funcții pentru monitorizarea utilizării sistemului şi aplicarea ajustărilor necesare.

Dintre cele mai utilizate **SGBD-uri**: Oracle, Access, MySQL, Sybase, FoxPro, Paradox, dBASE.

4. TIPURI DE BAZE DE DATE. MODELE DE DATE

Putem clasifica bazele de date astfel:

a) Dupa distribuția datelor:

a1) BD locale - toate fişierele care compun baza de date se găsesc pe un acelaşi sistem de calcul,

a2) BD distribuite - fişierele care compun baza de date sunt distribuite pe mai multe sisteme de calcul,

b) Dupa modelul de date (i.e. specificațiile de arhitectură a datelor). Modele clasice sunt**:**

b1) BD relaționale,

b2) BD ierarhice,

b3) BD în rețea.

Modelul de date este o reflectare a modelului conceptual, care reprezintă "lumea reala" prin concepte de tip entitate, relație şi atribute. O *entitate* este un anumit aspect al acestei "lumi reale", care are o existență independentă şi poate fi identificat în mod unic. Toate modelele clasice păstrează înregistrarea ca orientare fundamentală.

BAZE DE DATE RELAȚIONALE

Acestea sunt cele mai utilizate baze de date, pentru că sunt uşor de înțeles şi de manevrat. Structura lor logică este de tip tablou cu relații între linii şi coloane. Ne imaginăm că baza de date din figura I.6 conține informații privind pacienți purtători ai unei maladii ereditare, motiv pentru care s-au luat în observație şi descendenții acestora.

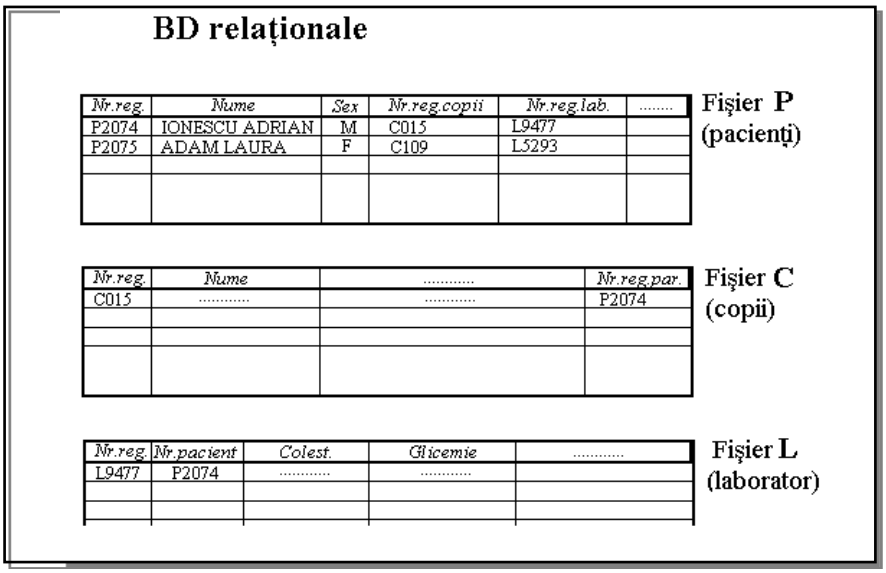


Figura I.6. Bazele de date relaționale sunt organizate ca nişte tabele cu relații între linii şi coloane (în acest exemplu numărul de registru este informația care face "conexiunea" între cele trei fişiere)

Căutarea în baza de date se face prin comparație între criteriile de căutare şi datele propriu-zise - valoarea de adevarat/fals, obținută ca rezultat, conducând la crearea unor înregistrări virtuale, care formează apoi tabele virtuale. Aceste tabele virtuale se obțin ca rezultat al căutării şi ele conțin răspunsul la întrebări de genul: "Care sunt rezultatele obținute la ultimele analize de laborator de pacientul IONESCU ADRIAN şi de copiii

acestuia?", sau "Există înregistrate datele privind evoluția părinților lui IONESCU ADRIAN? Dacă da, care sunt ele ?".

Aceste criterii de căutare se stabilesc de către utilizator, iar modificările în relațiile logice dintre câmpurile şi înregistrările bazei de date se definesc şi se modifică cu un efort minim. Bazele de date relaționale sunt foarte flexibile şi uşor expandabile.

BAZE DE DATE IERARHICE

Diagramele unor asemenea baze de date sunt arborescente: fiecare element este subordonat unui singur element de pe nivelul precedent al bazei de date şi numai unuia. Dependența unui segment de alte segmente de date de nivel superior se exprimă printr-un punctator (adresa), ceea ce conduce la o economie considerabilă de spațiu şi se simplifică regăsirea informațiilor de bază.

Prin această organizare eficientă timpul de prelucrare se scurtează substanțial. Plata pentru această eficiență este o suplețe mult mai scazută la schimbarea procedurii de prelucrare. Pentru volume mari de date şi prelucrări intense, avantajele oferite sunt însă determinante.

BAZE DE DATE ÎN REȚEA

Bazele de date în rețea sunt asemănătoare cu cele ierarhice, doar că un "copil" poate avea mai mult decât un singur "părinte", ceea ce le face mai flexibile dar şi mai puțin eficiente în operațiile de interogare.

MODELUL RELAȚIONAL DE REPREZENTARE ŞI REGĂSIRE A DATELOR. CARACTERISTICILE SGBD-urilor RELAȚIONALE

Există câteva familii mari de limbaje relaționale:

a) **Limbajele orientate pe transformări** - constituie o clasă de limbaje neprocedurale care, cu ajutorul relațiilor transformă datele de intrare în ieşirea dorită de utilizator. Aceste limbaje (cel mai cunsocut este SQL - *Structured Query Language*) produc structuri uşor de înțeles şi de manipulat în termeni practici: ce trebuie obținut, pornind de la ce este cunoscut (descriu doar modul în care datele sunt organizate şi pot fi regăsite).

Caracteristici: au nevoie de limbaj gazdă şi trebuie incluse în pachetele de aplicație.

SQL a cucerit piața atât datorită calităților sale, cât şi faptului că a fost standardizat de către ANSI (*American National Standards Institute*) şi a fost inițial promovat de către IBM.

b) **Limbaje bazate pe algebra relațională** - utilizează o serie de operatori algebrici relaționali (permutare, proiecție, restricție, selecție, împărțire, reuniune, intersecție, diferență, concatenare etc.).

Limbajul algebric relațional este un limbaj procedural complet, dar dificil pentru necunoscători. El se bazează pe utilizarea unui ansamblu de operatori cu ajutorul cărora se acționează asupra uneia sau mai multor relații din cadrul unei baze de date relaționale, drept rezultat obținându-se o nouă relație.

c) **Limbaje relaționale de tip grafic** - modul de lucru: utilizatorul completează o serie de răspunsuri, pe un exemplu, prin care sistemul "ghiceşte" ce trebuie făcut şi generează instrucțiuni corespunzătoare ale limbajului.

PROBLEME SOCIALE

Ca şi în cazul altor tehnologii, există aspecte variate şi uneori controversate legate de utilizarea bazelor de date electronice. Ele se asociază mai ales cu problemele de securitate a datelor şi cu faptul că se pot interoga colecții uriaşe de date aflate la distanțe mari, cu un efort minim.

Sunt şi cazuri în care apar probleme legate de dreptul de a colecta şi utiliza informația încă de la început sau dreptul de a da informația colectată către terțe părti. Multe dintre aceste probleme sunt încă fără un răspuns clar – exemple adaptate după [Brookshear 2007]:

- în ce masură poate o universitate face uz de datele despre studenții săi: (a) numele şi adresele; (b) notele? Poate face publică distribuția notelor fără a da numele?
- în ce măsură poate un spital să făcă uz de informațiile referitoare la pacienți - poate folosi datele pentru cercetare? Se schimbă ceva dacă datele sunt de-identificate? Le poate da unor instituții care fac cercetare farmaceutică?
- în Statele Unite există înregistrată informație ADN despre toți deținuții federali, iar baza poate fi consultată în situația unor investigații criminalistice – este etic ca această informație să fie utilizată şi pentru cercetare genetică?
- poate o bancă să dea informații referitoare la obiceiurile de cheltuieli ale clienților? Când, în ce condiții?

BIBLIOGRAFIE ŞI REFERINȚE

JH van Bemmel, MA Musen (eds). *Handbook of Medical Informatics*. Springer Verlag, Heidelberg, 1997

P Beynon-Davies. *Database systems* (2nd Edition). Macmillan Press, Houndmills UK, 2000

JG Brookshear. *Computer science: an overview* (9th Edition). Addison Wesley, Boston, 2007

E Coiera: *Guide to Medical Informatics, the Internet and telemedicine*. Chapman &Hall, London, 1997

EH Shortliffe, LE Perreault (eds). *Medical Informatics. Computer applications in healthcare and biomedicine* (2nd Edition). Springer Verlag, New York, 2001

T Spircu, S. Țigan: *Informatica în Medicină*. Ed. Teora, Bucureşti, 1995

Partea a II-a

BIOSTATISTICĂ

INTRODUCERE ÎN BIOSTATISTICĂ

Marea majoritate a cunoştinţelor manevrate în ştiinţele naturii, inclusiv cele medicale se bazează pe observaţii şi studii asupra mediului. Una dintre caracteristicile care frapează de la început este **variabilitatea.** Indivizii au diferite înălţimi, greutăţi, etc. Am observat însă cu toţii că variaţile observate sunt relativ limitate în intervale pe care le considerăm "rezonabile" sau "normale" iar ieşirea înafara intervalului reprezintă cel mai adesea ieşirea din sfera a ceea ce numim "normal". Privind astfel lucrurile am putea spune că ştiinţele medicale se ocupă cu depistarea acestor variaţii, cauzele şi metodele de revenire în domeniul valorilor normale. De fapt trebuie să se înceapă cu definirea limitelor în care încadrăm "normalul". Putem deja sesiza că acest lucru nu este deloc uşor fiindcă vom stabili aceste limite print-un studiu asupra unui grup de indivizi pe care îi considerăm "normali" încă înainte de a avea definit normalul. Vom mai observa că variabilitatea poate fi destul de ridicată; în plus, repetând studiul pe un alt grup obţinem alte limite, deci apar semne de îndoială privind stabilirea limitelor şi va fi firească întrebarea: "cum putem defini un interval rezonabil şi cât de mare încredere putem avea în compararea unei situaţii reale cu aceste date generale"? Acesta este rolul biostatisticii, care pe baza unei fundamentări matematice solide, în special teoria probabilităţilor, ne permite să ne orientăm printre datele atât de diverse ca cele oferite de viaţa de zi cu zi. Caracterul probabilist al interpretărilor este oarecum contrastant cu modelul "exact" impus de educaţia uzuală din matematică; de aceea se susţine că statistica nu este doar o ştiinţă, ci **un mod de gândire;** în matematică numerele 130 şi 135 sunt evident diferite; în gândirea statistică nu vom mai fi atât de siguri că sunt diferite (dacă ar fi vorba de exemplu de două măsurători de tensiuni arteriale sistolice ale unui individ vom ajunge probabil cel mai des la concluzia că nu sunt diferite!).

Ca orice ştiinţă, biostatistica operează cu câteva concepte de bază care vor fi prezentate în cele ce urmează.

1. INFERENŢA STATISTICĂ

1.1 CONCEPTELE DE BAZĂ ALE BIOSTATISTICII

A. Individ populaţie

Definiţie. **Individ** (element, unitate statistică) - concept de bază ce reprezintă forma individuală caracteristică fenomenului studiat şi supusă operaţiei de măsurare a unor parametri (mărimi).

Un individ este considerat bine definit dacă este:

- identificat concret
- localizat în timp (moment sau interval în care se consideră că nu se modifică sensibil caracteristicile studiate)
- localizat în spaţiu.

Exemplu: într-un studiu asupra dezvoltării copiilor, noţiunea de individ este asociată unui copil anume, fiind precizat şi momentul în care s-au efectuat măsurătorile precum şi localizarea spaţială a studiului.

Observaţie: "individ" nu este neapărat o persoană; într-un studiu făcut pe şobolani, individul va fi un şobolan, într-o probă de sânge va fi o hematie etc.
Definiţie: **Populaţie** (colectivitate statistică) reprezintă ansamblul tuturor indivizilor la care se referă studiul şi care au cel puţin o proprietate comună.

Populaţia este bine definită dacă este:
- localizată în timp
- localizată în spaţiu
- identificată caracteristica ce este comună indivizilor din populaţie.

Observaţii:
- numărul de indivizi dintr-o populaţie se numeşte **volumul** populaţiei
- populaţiile pot fi finite sau infinite
- indivizii îşi pierd individualitatea în interiorul unei populaţii.

Exemple de populaţii: - copiii în vârstă de 10 ani din judeţul Timiş, în anul 1995
- limfocitele T ale bolnavilor de hepatită B.

B. Obiectul şi metodele biostatisticii

Definiţie: Biostatistica este disciplina care îşi propune studiul caracteristicilor unei populaţii.

Metode de studiu pentru evaluarea caracteristicilor populaţiei**:**

- recensământ - metodă de determinare exactă a caracteristicilor populaţiei; localizarea în timp este restrânsă la un moment; tot foarte bine precizată este şi delimitarea spaţială. Recensământul este o operaţie laborioasă şi foarte costisitoare fiind utilizat rar, pentru culegerea unor date exacte strict necesare. Deseori în practică nu este necesară precizia oferită de recensământ, fiind suficiente date aproximative, studiile devenind mult mai ieftine.

- screenning - metodă asemănătoare recensământului utilizată de obicei pentru depistarea în cadrul unei populaţii a indivizilor având o abatere deosebită a unui parametru (depistarea precoce a unor afecţiuni grave sau cu consecinţe deosebite); nu este necesară localizarea în timp cu strictеţea recensământului. Fiind o operaţie destul de costisitoare, eficienţa creşte prin alegerea unei selecţii din populaţie conform unor factori de risc; există o întreagă metodologie pentru optimizarea screeningurilor.

- selecţie (eşantionare) - metoda cel mai des folosită oferind rezultate cu precizie satisfacătoare şi un cost mult redus; pentru studiu se alege din populaţie o submulţime numită **eşantion** (lot, grup), măsurătorile fiind efectuate numai pe indivizii eşantionului studiat.

C. Inferenţa statistică

Rezultatele obţinute pe un eşantion le vom considera valabile pentru întreaga populaţie.

Definiţie: Operaţia de generalizare a caracteristicilor unui eşantion la nivelul întregii populaţii se numeşte **inferenţă statistică.**

Inferenţa statistică este operaţia fundamentală a statisticii şi în jurul ei gravitează majoritatea aspectelor teoretice. Importanţa acestei operaţii poate fi mai bine sesizată dacă ne gândim că eşantionul poate să reprezinte un procent infim din întreaga populaţie (să zicem 1:1000); putem foarte uşor aluneca spre concluzii eronate dacă eşantionul nu reprezintă întreaga diversitate pe care o întâlnim în populaţie. Să ne oprim deci puţin asupra operaţiei de selecţie în eşantion (teoria selecţiei).

D. Eşantion reprezentativ

Definiţie: Eşantionul care conţine proporţional indivizi reprezentând toate caracteristicile populaţiei poartă numele de **eşantion reprezentativ.**

Criterii pentru eşantionul reprezentativ

O serie de concluzii interesante pentru statistică s-au putut desprinde cu ocazia sondajelor efectuate în peajma alegerilor, partidele politice fiind deosebit de interesate de aceste rezultate. Instituţiile care se ocupă cu efectuarea acestor sondaje utilizează de obicei eşantioane reprezentând - 1:1000 - 1:100 din populaţie. Este des pomenită în cărţile de statistică o întâmplare cu ocazia alegerilor din SUA din anul 1936. În sondajul efectuat de revista "The Literary Digest" cu puţin înainte de alegeri, candidatul Alfred London părea să aibă un avantaj sensibil, însă alegerile au fost câştigate detaşat de F. D. Roosvelt, deşi erorile sunt neaşteptat de mari, dacă ne gândim că sondajul a fost efectuat pe cca 10 milioane de alegători (din cca 40). Explicaţia constă însă tocmai în alegerea defectuoasă a eşantionului: fiecare al treilea alegător înregistrat din Chicago, nume alese la întâmplare din listele diferitelor cluburi, din cartea de telefon, etc., "favorizând" selecţia din mediul urban faţă de rural, din cei cu venituri mari faţă de cei cu venituri mici, dintre bărbaţi faţă de femei. În plus, sondajul s-a efectuat parţial prin poştă şi numai un sfert din cei solicitaţi au răspuns. S-au putut stabili ulterior reguli pentru alegerea eşantioanelor reprezentative care contează mai mult decât mărimea eşantionului.

Criteriile *pentru selecţia în eşantionul reprezentativ:*

Echiprobabilitate: toţi indivizii din populaţie să aibă aceeaşi probabilitate de a fi aleşi în eşantion.

Independenţă: alegerea unui individ să fie independentă de alegerea altui individ în eşantion.

E. Metode de selecţie

a. Sondajul simplu. Populaţia de studiat este considerată omogenă şi fiecare individ este ales în mod aleator (întâmplător) din întreaga populaţie. Respectarea practică a criteriilor enumerate anterior este dificilă datorită mai multor factori: subiectivitatea celui ce face eşantionarea, necooperarea unor indivizi selectaţi, unele condiţii tehnice (când "indivizii" sunt animale, celule etc.). De aceea selecţia se face, pentru populaţii finite, prin cartografierea populaţiei (numerotarea indivizilor luaţi în evidenţă) şi apoi generarea de numere aleatoare (sau folosirea unor tabele de numere aleatoare) care precizează indivizii selectaţi.

b. Sondajul dirijat. Populaţia este deseori heterogenă şi poate fi divizată în mai multe categorii, numite "straturi" în teoria selecţiei. (De exemplu un eşantion reprezentativ pentru populaţia ţării noastre trebuie să cuprindă proporţiile cuvenite de bărbaţi respectiv femei, de persoane din mediul rural sau urban, din diferite regiuni ale ţării). În sondajul dirijat se vor selecta în eşantion un număr de indivizi din fiecare strat, proporţional cu ponderea stratului în populaţie. În interiorul fiecărui strat se aplică regulile sondajului simplu. Există mai multe variante pentru punerea în practică, cel mai adesea, (în anchetele stării se sănătate) alegându-se din fiecare strat localităţi şi eşantionare aleatoare (gospodării, nr. indivizi) din localităţile alese.

c. Sondajul mixt. Versiune care îmbină sondajul dirijat cu cel simplu.

Trebuie menţionat aici că există o serie de reguli (ce vor fi în parte discutate ulterior) prin care se stabileşte **numărul minim** de indivizi din eşantion pentru studiul propus.

F. Tipuri de eşantioane

Deşi dorința noastră, în cursul selecției eşantionului, este cel mai adesea cea de a evita orice factori care ar influența echiprobabilitatea, constatăm că, practic, numeroşi factori au tendințe, uneori puternice, de a afecta acest criteriu fundamental. De exemplu dacă dorim să selectăm pentru o experință un lot de şoareci din biobază (crescătorie), este posibil să apară factori subiectivi (îngrijitoarele au uneori "simpatii" pentru animalele îngrijite - unii şoareci sunt mai drăguți, cu mustăți mai lungi etc. şi încearcă amânarea unui eventual destin tragic) sau obiectivi (dacă alegerea unor şoareci la întâmplare dintr-o cuşcă înseamnă a lua pe cei pe care îi poți prinde deja putem sesiza că îi vom prinde pe cei mai puțin abili în a evita capturarea lor - cu alte cuvinte alegerea nu mai este chiar întâmplătoare; în această categorie intră şi sondajul prin poştă: serozitatea de a răspunde la un sondaj nu este egal distribuită în toate straturile!).
Definiție: un factor (tendință) care influențează probabilitatea de selecție a unui individ într-un eşantion se numeşte **bias.**

În funcție de prezența sau absența unui astfel de factor în procesul de selecție putem distinge:

- eşantioane neselective ("unbiased") - care respectă echiprobabilitatea
- eşantioane selective ("biased") în care, cu sau fără ştirea noastră, un factor a influențat componența lotului. În marea majoritate a cazurilor eşantioanele selective sunt evitate; excepție fac studiile din "analiza riscului" unde aceşti factori sunt chiar căutați pentru eventuala definire a unor "straturi".

1.2. VARIABILE

Studiile experimentate, indiferent de natura lor, se concretizează prin culegerea unor date.

A. Definiție

Mărimile asupra cărora este orientat un studiu şi se culeg date poartă numele de **variabile** sau **caracteristici.**

B. Tipuri de variabile

a. Variabile numerice (se mai numesc **cantitative** sau **cardinale)** sunt cele ale căror valori sunt exprimate prin numere, pornind de la o **unitate de măsură** bine definită. Valoarea numerică propriu-zisă depinde de unitatea de măsură şi de precizia instrumentului de măsură.

Exemple de variabile numerice:
- înălțimea unui individ (talia) se exprimă de obicei în cm sau m (în picioare şi inch în unele țări anglo-saxone);
- greutatea - de obicei în kg cu precizia de 1 kg pentru adulți şi 10g-100g pentru copii (în "pounds" şi "ounces" în unele tări anglo-saxone);
- frecvența cardiacă - se exprimă în bătăi/minut;
- pH-ul sanguin - se exprimă în unități pH, etc.

Variabilele numerice pot fi de două tipuri:
- scară proporțională: valoarea zero-originea- este aceeaşi indiferent de unitatea de măsură
- scară de intervale: la schimbarea unității de măsură rămân proporționale numai intervalele (ex: temperatura în ^{0}C şi ^{0}F).

Variabilele cantitative se mai pot împărți în:
- variabile **continue** - exprimate prin numere reale (ex: pH-ul sanguin)

- variabile **discrete** - exprimate prin numere întregi sau raționale, având numai anumite valori posibile (ex: frecvența cardiacă).

În practică, datorită preciziei limitate a instrumentelor cu care se efectuează măsurările, putem aborda toate variabilele ca discrete; deseori nici nu este necesară împingerea precizie măsurătorilor la maximul tehnic posibil, fiind suficientă o precizie limitată, specificată după necesități (ex: greutatea unui individ)**.**

b. Variabile ordinale (se mai numesc variabile **rang)** - sunt exprimate prin numere conform unei **scări** (scale) convenționale care acceptă relația de ordine (mai mare, mai mic, egal) sau ataşează eventual valori numerice conform unor criterii convenționale. Specific variabilelor ordinale este exprimarea prin numere, dar absența unei unități de măsură.

Exemple de variabile ordinale:

- "nota" obținută de un individ la un examen indiferent de forma de examinare; nota nu este altceva decât reflectarea sub formă de număr a poziției pe o scară cu trepte convenționale: nu se poate niciodată afirma că "diferența" (distanța) între 9 şi 10 "este egală" cu diferența (distanța) între 5 şi 6!
- "ierarhia" în scara Luscher de preferință a culorilor dintr-o paletă de 6 (sau 8) culori date
- gradul de apreciere al efortului pe scara Berg (0-20 de la "f.f.uşor" până la "epuizant"); etc. **c. Variabile calitative (**se mai numesc **nominale) -** sunt exprimate prin nume sau simboluri ce definesc diferite "clase" de calități; între calitățile claselor nu se poate (în general) stabili o relație de ordine; cele mai frecvente sunt calitățile definite dihotomizant astfel încât să apară numai două clase; teoria demonstrează că toate celelalte cazuri se pot reduce în final la abordarea dihotomică (cu două clase); variabilele cu numai două valori posibile se mai numesc şI variabile **alternative.**

Exemple de variabile calitative:

- grupa sanguină: O/A/B/AB - patru valori posibile
- sexul: M/F - două valori posibile
- starea pacientului după tratament: ameliorat/neameliorat - două valori posibile, etc.

C. Caracteristicile variabilelor

Pentru un individ (obiect) valoarea asociată variabilei este definită pentru un moment dat. De obicei culegem un ansamblu de valori corespunzând indivizilor din eşantion urmând ca, prin inferență statistică, sa încercăm să caracterizăm **starea** populației la momentul respectiv.

În cazul în care urmărim şi evoluția în timp a valorilor variabilei studiate obținem o **serie temporală** care ne dă o reprezentare dinamică, în timp, a mărimii studiate. Putem întâlni serii de momente sau serii de intervale.

Uneori studiul se îndreaptă spre sesizarea unor diferențe în funcție de repartiția în spațiu a valorilor individuale; în acest caz obținem o **serie spațială.**

D. Mărimi deterministe şi aleatoare

a. Mărime deterministă - este o marime a cărei valoare este bine definită la un moment dat şi care prin repetarea măsurării ne aşteaptăm să obținem din nou aceeaşi valoare. Eventualele variații se pot datora numai operației propriu-zise de măsurare.

Exemple: lungimea unui obiect , concentrația unei soluții etc.

b. Mărime aleatoare (sau **statistică)** - este o mărime a cărei valoare nu se repetă cu necesitate prin repetarea măsurătorii, chiar dacă încercăm să păstrăm nemodificate condițiile experimentale.

Exemple: rezultatul la aruncarea zarului (6 valori posibile) sau numărul de dezintegrări/secundă al unei surse radioactive (cu variații într-un anumit interval), structura genomului unui descendent (cu câteva variante posibile). Procesele care generează marimi statistice se mai numesc şi **stochastice.**

2. PARAMETRII STATISTICI

În biostatistică se utilizează frecvent introducerea unor noțiuni pornind de la exemple, metodă pe care o vom utiliza şi noi în continuare.

Exemplul 1: Efectuăm un studiu privind dezvoltarea copiilor din Timişoara în 1995. Luăm un eşantion de 324 copii în vârstă de 10 ani din Timişoara obținând valorile din tabelul de mai jos.

Tabelul 1. Datele cu înălțimile unui grup de copii

Înălțime	Frecvența		Înălțime	Frecvența		Înălțime	Frecvența
126	1		134	25		142	13
127	1		135	28		143	12
128	3		136	34		144	8
129	4		137	37		145	3
130	7		138	30		146	2
131	12		139	31		147	2
132	11		140	22		148	0
133	19		141	18		149	1

Datele din acest tabel pot fi reprezentate grafic sub forma unei histograme (fig. II.1)

Analizând aceste date, putem observa că:
- valorile foarte mici sau foarte mari (extremele), sunt rare, majoritatea valorilor fiind situate în zona centrală; acest aspect va fi surprins de unii parametri caracteristici numiți "indicatori ai tendinței centrale"
- variațiile în jurul parametrului tendinței centrale pot fi mai mari sau mai mici (valorile individuale pot fi mai grupate sau mai împrăştiate); acest aspect va fi surprins de alți parametri numiți "indicatori de dispersie".

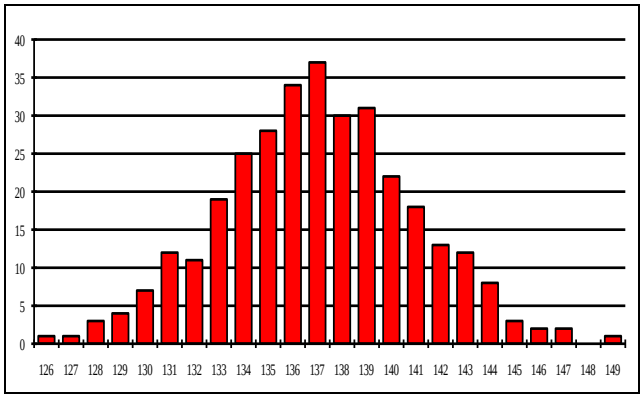


Figura II.1. Histograma înălțimii unui grup de copii

2.1. INDICATORII TENDINȚEI CENTRALE

În funcție de tipul variabilei se recomandă folosirea unor diverşi indicatori ai tendinței centrale.

A. Media aritmetică

Este cel mai folosit indicator al tendinței centrale. Dacă avem un eşantion format din *N* indivizi şi notăm valorile variabilei studiate cu X_i, *i=1,....,**N*** (citim "indicele *i* luând valori de la 1 la **N"**), atunci media aritmetică a variabilei *X*, notată cu $\overline{X}$ este dată de relația:

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i = \frac{X_1 + X_2 + \ldots + X_N}{N} \qquad \text{(II.2.1.a)}$$

1. În cazul eşantioanelor mai mari, anumite valori pot să apară de mai multe ori (ca de ex. în tabelul II.1); dacă notăm frecvența absolută de apariție a fiecărei valori x_j cu n_j, atunci media aritmetică se mai numeşte **medie ponderată** şi este dată de relația:

$$\overline{X} = \frac{1}{N}\sum_{j=1}^{K} n_j X_j \qquad \text{(II.2.1.b)}$$

unde *K* reprezintă numărul de clase, iar frecvențele respectă relația:

$$N = \sum_{i=1}^{K} n_j \qquad \text{(II.2.1.c)}$$

Observație: În cazul în care o "clasă" *j* nu conține numai indivizi care au exact aceeaşi valoare X_j ci apar variții (deci "clasa" reprezintă de fapt un "strat", atunci definim mai întâi media stratului j:

$$\overline{X}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} X_{ij} \qquad \text{(II.2.1.d)}$$

şi media pe întregul eşantion va fi:

$$\overline{X} = \frac{1}{N}\sum_{j=1}^{K}\sum_{i=1}^{n_j} X_{ij} \qquad \text{(II.2.1.e)}$$

2. În cazul unei variabile numerice continue, în loc de frecvența clasei folosim valoarea funcției de distribuție, deci media aritmetică va fi definită prin:

$$\overline{X} = \frac{1}{X_{\max} - X_{\min}} \int_{X_{\min}}^{X_{\max}} f(x)\, dx \qquad \text{(II.2.1.f)}$$

B. Mediana

În cazul variabilelor ordinale este mai potrivit a folosi în loc de media aritmetică un alt indicator, mediana definită astfel:

Definiție: Mediana este un indicator statistic al tendinței centrale care împarte lotul în două părți egale; 50% din indivizi au valori mai mici decât mediana, 50% au valori mai mari.

Pentru calculul propriuzis al medianei se procedează astfel:
- se formează un şir ordonat crescător al tuturor celor *N* valori
- dacă *N* este impar (adică *N=2p+1),* atunci mediana

$$M_e = X_{p+1} \qquad \text{(II.2.2.a)}$$

Valoarea calculată cu (II.2.2.a) este considerată exactă dacă :

$$X_p \langle X_{p+1} \langle X_{p+2} \qquad \text{(II.2.2.a')}$$

- dacă *N* este par (adică *N = 2p*), atunci mediana M_e este aproximată de

$$M_e = \frac{X_p + X_{p+1}}{2} \qquad \text{(II.2.2.b)}$$

Valoarea dată de (II.2.2.b) este considerată exactă dacă:

$$X_{p-1} \le X_p \langle X_{p+1} \le X_{p+2} \qquad \text{(II.2.2.b')}$$

sau

$$X_{p-1} \langle X_p = X_{p+1} \langle X_{p+2} \qquad \text{(II.2.2.b")}$$

Dacă nu sunt îndeplinite condițiile (II.2.2.a'), (II.2.2.b') sau (II.2.2.b'') atunci valoarea care se repetă de mai multe ori (de $\boldsymbol{n_{im\,i}}$) defineşte un "interval median" de lățime $\boldsymbol{h_{im}}$; mai notăm frecvența cumulată până la intervalul median (până la limita inferioară inclusiv) cu $\boldsymbol{f_{im}}$; în acest caz mediana poate fi aproximativă cu relația:

$$M_e = X_{im} + \frac{h_{im}}{n_{im}}\left(\frac{N}{2} - f_{im}\right) \qquad \text{(II.2.2.c)}$$

unde:

$$f_{im} = \sum_{i=1}^{im-1} n_i \qquad \text{(II.2.2.c')}$$

Observație: Deşi nu este recomandabil, destul de des apare calculată media aritmetică (fiind mai simplu de calculat) în loc de mediană în cazul variabilelor ordinale. În cazul distribuțiilor simetrice media aritmetică şi mediana coincid; totuşi pentru distribuții asimetrice ele au valori diferite!

C. Moda (modul)

În cazul variabilelor calitative (nominale), media aritmetică sau mediana nu au sens; indicatorul tendinței centrale utilizabil se numeşte **modă** sau **mod.**

Definiție: Moda (Mo) reprezintă cea mai frecventă valoare.

Moda poate fi definită şi pentru variabilele numerice sau ordinale. Pentru exemplul din tabelul II.1. moda este de 137 cm. În cazul distribuțiilor simetrice pentru variabilele numerice media aritmetică, mediana şi moda coincid.

Poziția relativă a modei, medianei şi mediei aritmetice pentru distribuții asimetrice este vizibilă în figura II.2.

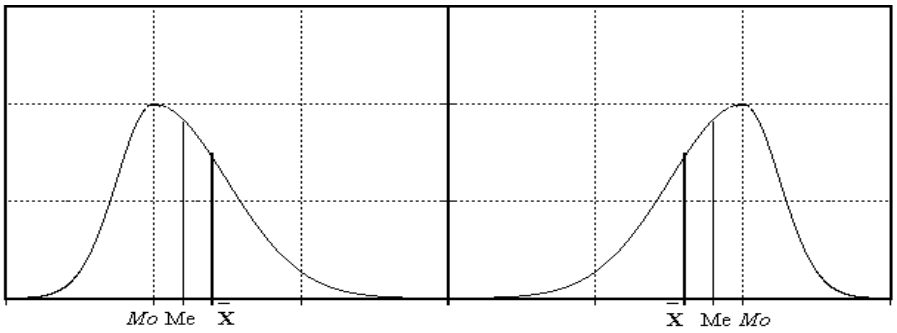


Figura II.2. Moda, mediana şi media aritmetică pentru distribuții asimetrice

Pentru variabile numerice, dacă obținem o curbă de distribuție cu un singur maxim, ea se numeşte **unimodală;** în cazul în care are două maxime (chiar dacă diferite), distribuția se numeşte **bimodala** (fig. II.3). Similar, pentru mai multe maxime putem întâlni distribuții **multimodale.** În cazul populațiilor omogene ne aşteptăm doar la distribuții unimodale. Depistarea unei distribuții bi sau multimodale este cel mai adesea un indiciu al unei populații neomogene din care s-a extras eşantionul, fiind cel mai probabil o suprapunere a două populații cu caracteristici diferite. Distribuțiile bi- sau multi-modale merită un studiu mai amănunțit.

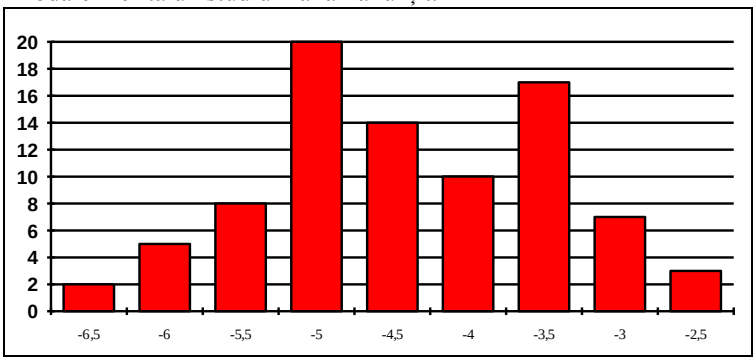


Figura II.3. Distribuția bimodală: distribuția pragului de sensibilitate pentru detecția gustului "amar" c=concentrația de chinină la care este sesizat gustul amar)

Observație: În cazul variabilelor numerice alura graficului de distribuție depinde puternic de modul în care definesc clasele (lățimea intervalelor). Deseori autorii sunt tentați a lua lățimea clasei egală cu precizia cu care s-a efectuat măsurarea (de ex. pentru înălțimea copiilor prezentată în tabelul II.1. precizia măsurătorii a fost de 1 cm şi de aceea în figura II.1. s-a reprezentat distribuția conform acestei precizii). În biostatistică se recomandă ca numărul de clase utilizat să nu fie prea mare, astfel încât fiecare clasă să fie destul de reprezentată. Dacă se notează cu X_m şi X_M valorile extreme găsite (minimă, respectiv maximă), si cu N numărul total de indivizi din eşantion, atunci lățimea unei clase pentru "histogramă" poate fi aproximată prin relația:

$$h = \frac{X_M - X_m}{1 + 3{,}322 \lg N} \qquad \text{(II.2.3)}$$

(Pentru datele din tabelul II.1 obținem h ≈ 2,8 ≈ 3 cm deci reprezentarea recomandabilă ar avea clasele de înălțime ale copiilor de câte 3 cm: 124-126, 127-129, 130-132, 133-135, 136-138, 139-141, 142-144, 145-147, 148-150 în total 9 clase în loc de 25).

D. Alți indicatori ai tendinței centrale

Foarte rar este posibil a întâlni şi alți indicatori ai tendinței centrale:

- media geometrică: $$\overline{X}_g = \left(\prod_{i=1}^{N} \mathrm{X}_i \right)^{1/N} \qquad \text{(II.2.4.a)}$$

- media armonică: $$\frac{N}{\overline{X}_h} = \sum_{i=1}^{N} \frac{1}{\mathrm{X}_i} \qquad \text{(II.2.4.b)}$$

2.2. INDICATORI DE DISPERSIE

Gradul de variabilitate al valorilor individuale într-o populație, vizibilă într-un eşantion se exprimă printr-un set de parametri statistici numiți "indicatori de dispersie". Există mai multe posibilități de a exprima acest grad de variabilitate:

A. Domeniul de valori (amplitudinea; engl. range)

Este un indicator simplu, furnizând doar informații asupra ordinului de mărime al variabilității.

$$R = X_{max} - X_{min} \qquad \text{(II.2.5.a)}$$

unde X_{min} şi X_{max} reprezintă valoarea absolută minimă, respectiv maximă ale variabilei analizate.

Se foloseşte mai rar, de obicei în prezentarea părții introductive a studiului, limitele vârstelor subiecților dintr-un lot sunt adesea prezentate în acest mod.

B. Abatarea centrală. Eroarea medie absolută

Definiție: Distanța unei valori individuale față de valoarea medie se numeşte abatere centrală:

$$\varepsilon_i = X_i - \overline{X}$$

Abaterile centrale pot fi pozitive sau negative. Ele au propietatea evidentă că:

$$\sum \varepsilon_i = 0$$

De aceea, pentru a caracteriza gradul de variabilitate, se folososesc valorile absolute ale abaterilor centrale. O mărime ce poate fi folosită ca măsură a variabilității este:

$$\varepsilon_a = m = \frac{1}{N}\sum|\varepsilon_i| = \frac{1}{N}\sum|X_i - \overline{X}| \qquad \text{(II.2.5.c)}$$

şi se numeşte **eroare medie absolută**.

C. Deviația standard

Amplitudinea, definită anterior, nu ne spune nimic cu privire la repartiția indivizilor între minim şi maxim. Informații mai complete primim dacă urmărim şi această repartiție.

În figura II.1. am prezentat histograma înălțimii pe un grup finit şi destul de mic (cca 300), având o precizie oarecum redusă de exprimare a variabilei (precizia de 1 cm).

Ne putem imagina că, dacă am efectua masurări cu precizie mult mai mare, pe un lot foarte numeros, am putea obține curba reală de distribuție a înălțimilor într-o populație de copii de vârsta dată. S-a constatat că foarte multe mărimi observabile în natură se pot reprezenta print-o curbă de distribuție simetrică față de valoarea medie, sub forma unui "clopot", numit "clopotul lui Gauss" sau "curba distribuției normale (fig. II.4).

Ecuația curbei lui Gauss este:

$$f(x) = y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \text{(II.2.6.a)}$$

Observăm că în ecuație apar 2 parametri: μ şi σ.

. μ este indicatorul tendinței centrale, reprezintă "media" şi este valoarea în jurul căreia curba este simetrică;

. σ este indicatorul de dispersie, se numeşte "deviație standard" sau "abatere standard" şi arată gradul de împrăştiere a curbei în jurul mediei.

Deviația standard în curba lui Gauss permite urmărirea repartiției valorilor individuale în jurul valorii medii conform fig. II.4.

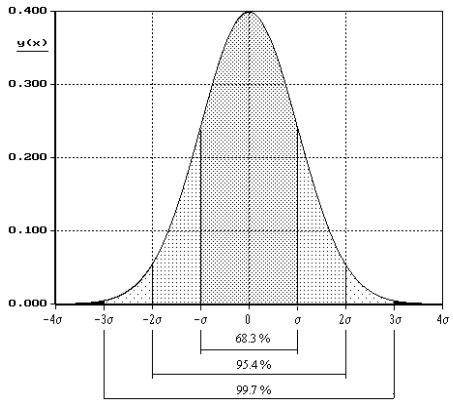


Figura II.4. Distribuția normal

Analizând figura putem spune că valorile individuale X_i se vor găsi în intervalele:

$X_i \in (\mu - \sigma, \mu + \sigma)$ în 68% din cazuri

$X_i \in (\mu - 2\sigma, \mu + 2\sigma$ în 95,4% din cazuri (II.2.6.b)

$X_i \in (\mu - 3\sigma, \mu + 3\sigma)$ în 99,7% din cazuri

În cazul lucrului pe un eşantion, în loc de media populației μ se va folosi media eşantionului, $\overline{X}$, iar în loc de deviația standard a populației se va folosi deviația standard a eşantionului (abatere standard) care se calculează cu relația:

$$s = \sqrt{\frac{\sum \varepsilon_i^2}{n-1}} = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n-1}} \qquad \text{(II.2.7.a)}$$

Exemplu: în urma unui studiu pe un eşantion format din n = 25 copii de 10 ani, în care am găsit înălțimea medie $\overline{X}$ = 137 cm şi deviația standard *S* = 5 cm putem afirma că cca 68% din copiii de 10 ani au înălțimea cuprinsă între 132 cm ($\overline{X} - S$) şi 142 cm ($\overline{X} + S$), cu alte cuvinte, probabilitatea ca înălțimea unui copil să fie între 132 - 142 cm este 68%.

În general, vom scrie astfel:

$X_i \in (\overline{X} - S,\ \overline{X} + S)$ cu p = 68,3%

$X_i \in (\overline{X} - 2S,\ \overline{X} + 2S)$ cu p = 95,4% (II.2.6.c)

$X_i \in (\overline{X} - 3S,\ \overline{X} + 3S)$ cu p = 99,7%

Definiție: putem acum defini deviația standard: S reprezintă gradul de variație a valorilor individuale în jurul mediei eşantionului.

Mărimea

$$S^2 = \frac{1}{n-1} \sum (X_i - \overline{X})^2 \qquad \text{(II.2.7.b)}$$

se numeşte **dispersie,** sau **abatere patratică** sau **varianță.**

Deseori se raportează indicatorul de dispersie la valoarea medie obținând o nouă mărime numită **coeficient de variație** exprimat în procente prin:

$$\text{C.V.} = 100 * S / \overline{X} \qquad \text{(II.2.7.c)}$$

Fiind o mărime relativă, se pot compara cu ajutorul ei serii având valori cu ordine de mărimi diferite.

Pentru loturi foarte mari, în formulele (II.2.7.a) şi (II.2.7.b) se foloseşte *n* în loc de *n-1*.

D. Eroarea standard a mediei

Scopul principal al unui studiu statistic este carcterizarea populației, nu a eşantionului. Din exemplul folosit până acum, am putea oare răspunde la întrebarea "Care este înălțimea medie a copiilor de 10 ani din Timişoara?" Este mare tentația de a răspunde: 137 cm (media eşantionului). Nu se poate însă să nu ne dăm seama că este

foarte posibil ca, repetând măsurătorile, pe un alt eşantion, să obţinem altă valoare medie, de exemplu 135,8; pe un al treilea eşantion 137,6 şi aşa mai departe. Este vreuna din aceste valori mai demnă de încredere decât celelalte? Nicidecum! Valoarea adevărată a mediei populaţiei, μ, se poate obţine experimental numai făcând măsurători pe întreaga populaţie. Concluzia pare demoralizantă la prima vedere. Totuşi, adesea nu este necesară cunoşterea foarte exactă a unui parametru, fiind suficientă încadrarea lui într-un interval suficient de îngust. Să vedem care sunt căile pentru estimarea acestui interval.

Presupunem că din populaţia studiată facem măsurători pe toate eşantioanele posibile de aceeaşi dimensiune ***n***, obţinând mediile $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_j, \ldots, \overline{X}_T$. Vom avea evident media populaţiei:

$$\mu = \frac{1}{T}\sum_{j=1}^{T} \overline{X}_j \qquad \text{(II.2.1.g)}$$

Analizând distribuţia acestor medii ale eşantioanelor vom observa (fig. II.5.a) că şi ele se aranjează aproximativ după o curbă Gauss, (dacă eşantioanele sunt destul de mari, n > 30) având faţă de curba din fig. II.4. două deosebiri:

a. variaţiile mediilor eşantioanelor se întind pe un interval mult mai îngust decât variaţiile valorilor individuale;

b. valoarea în jurul căreia sunt simetrice variaţiile este media populaţiei.

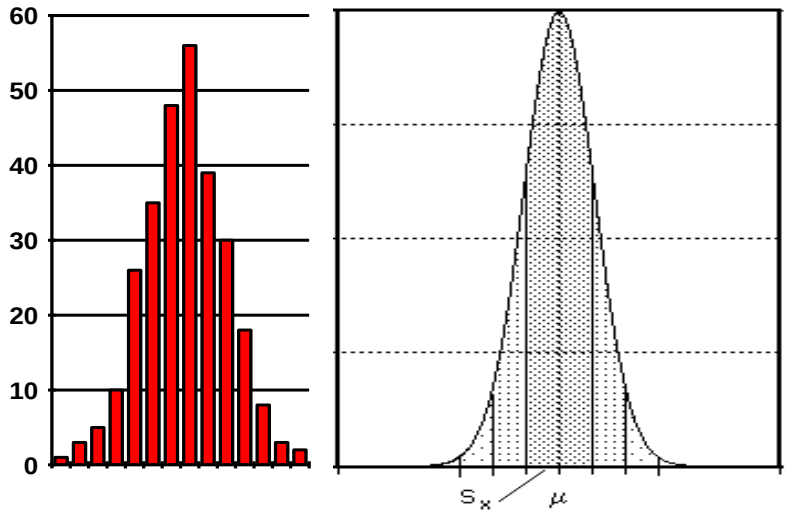


Figura II.5. Distribuţia mediilor eşantioanelor

Distribuţia mediilor eşantioanelor este caracterizată prin parametrii μ = media populaţiei şi $\sigma_{\overline{X}}$ = eroarea standard a mediei dată de formula:

$$\sigma_{\overline{X}} = \sigma / \sqrt{N} \qquad \text{(II.2.8.a)}$$

unde N = volumul populaţiei şi σ = deviaţia standard.

Să aplicăm formula (II.2.8.a) care pentru exemplul nostru concret devine:

$$S_{\overline{X}} = \frac{S}{\sqrt{n}} \qquad \text{(II.2.8.b)}$$

pentru eşantioane foarte mari, iar pentru eşantioane mici folosim

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} \qquad \text{(II.2.8.c)}$$

unde n = mr. indivizi din eşantion, N = volumul populației.

Deci înlocuind datele din exemplu avem:

$$S_{\bar{X}} = 5/\sqrt{25} \approx 1cm$$

Conform analizei schematizate în figura II.5.b putem afirma că 68% din mediile eşantioanelor de câte 25 copii vor avea media cuprinsă în intervalul (137 - 1, 137 + 1) adică între 136 şi 138 cm, sau în alte cuvinte, probabilitatea ca media unui eşantion oarecare să fie cuprinsă între 136 - 138 cm este 68%; Evident, nu pretindem că ştim cu exactitate media populației, μ , dar avem deja o localizare satisfăcătoare a sa, având posibilitatea de 68% de a fi încadrată în intervalul 136 - 138 cm.

Generalizând, putem scrie:

$$\mu \in (\bar{X} - S_{\bar{X}}, \bar{X} + S_{\bar{X}}) \quad \text{cu p = 68,3\%}$$

$$\mu \in (\bar{X} - 2S_{\bar{X}}, \bar{X} + 2S_{\bar{X}}) \quad \text{cu p = 95,4\%} \qquad \text{(II.2.8.d)}$$

$$\mu \in (\bar{X} - 3S_{\bar{X}}, \bar{X} + 3S_{\bar{X}}) \quad \text{cu p = 99,7\%}$$

Definiție: Eroarea standard a mediei: $S_{\bar{X}}$ reprezintă gradul de variație al mediilor eşantioanelor în jurul mediei populației.

Cu alte cuvinte, chiar dacă printr-un studiu pe un eşantion nu putem preciza cu exactitate parametrii caracteristici ai populației, putem totuşi să îi localizăm în anumite intervale, operație care se numeşte "estimare" şi care va fi analizată detaliat ulterior.

Să observăm însă că avem o relație de inversă proporționalitate între posibilitatea încadrării într-un interval şi lățimea intervalului: cu cât suntem mai siguri pe localizare, cu atât intervalul este mai larg. De aceea trebuie să găsim un compromis între siguranța localizării şi lățimea intervalului. Experința arată că o localizare cu precizie de 95% este satisfăcătoare din ambele puncte de vedere şi vom accepta această valoare pe tot parcursul cursului nostru.

E. Indicatori de dispersie ai variabilelor ordinale

Indicatorii de dispersie descrişi anterior, deviația standard şi eroarea standard a mediei sunt folosiți în special pentru variabilele cantitative propriu-zise. În cazul variabilelor ordinale, deşi se poate folosi şi deviația standard (şi eroarea standard a mediei), se recomandă folosirea unor indicatori specifici. Pentru a înțelege aceşti indicatori de dispersie vom introduce mai întâi alți parametri:

a. Cuantile. Indicatorul tendinței centrale specifice variabilelor ordinale, mediana, era definită ca valoarea care împarte lotul în două părți egale.

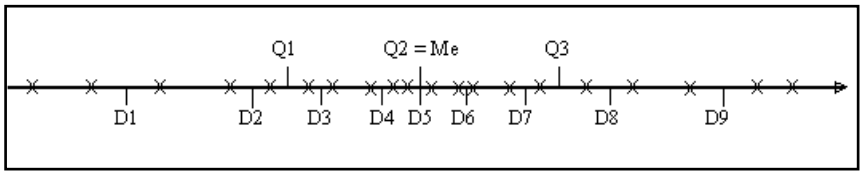


Figura II.6. Mediana, Cuartile, Decile pentru N=20

Putem, prin analogie, defini diverse alte mărimi numite **cuantile:** ca fiind valorile care împart lotul în **n** subclase echinumerice. Denumirile lor sunt prezentate în tabelul II.2.

Tabel II.2. Cuantile uzuale

Nr. clase	Simbolul valorilor	Denumire	Observații
2	Me	mediana	
4	Q1, Q2, Q3,	cuartile	Q2 = Me
10	D1, D2, ..., D9	decile	D5 = Me
100	C1,C2,... C99	centile	C50=Me C10 = D1, etc.
1000	P4, M2, ...,M999	promile	M10 = C1...

b. amplitudinea intercuatile (variație intercuatile):

$$Q_{ed} = (Q_3 - Q_1)/2 \qquad \text{(II.2.5.b)}$$

este o măsură a variabilității, valorile mai mari exprimând o variabilitate mai mare

c. coeficientul de variație inter-cuartil:

$$C.Q.= \frac{Q_d}{M_e} = \frac{Q_3 - Q_1}{Q_3 + Q_1} \qquad \text{(II.2.5.c)}$$

joacă rolul coeficientului de variație dat de (II.2.7.c) şi are valori între -1 şi +1.

F. Indicatori de dispersie ai variabilelor nominale

Pentru variabilele nominale indicatorul preferat al tendinței centrale este **moda**; fiecare calsa **i** este caracterizată prin procentul din eşantionul de volum **n:**

$$p_i = 100.\frac{n_i}{n} \qquad \text{(II.2.9.a)}$$

unde n_i este frecvența absolută a clase i.

Deviația standard a procentului este dată de relația:

$$S_p = \sqrt{\frac{p_i \cdot q_i}{n}} \qquad \text{(II.2.9.b)}$$

unde: $q_i = 100 - p_i$ (II.2.9.c)

În cazul unei populații finite de volum N

$$S_p = \sqrt{\frac{p_i q_i}{n}} * \sqrt{\frac{N-n}{N-1}} \qquad \text{(II.2.9.d)}$$

Pentru eşantioane suficient de mari, procentul în eşantion are distribuție normală şi permite interpretări similare cu cele prezentate anterior.

2.3. MEDII DE PUTERI: MOMENTE. MOMENTE CENTRATE

Abordarea teoretică a parametrilor statistici caracteristici unui set de valori permite generalizarea unor relații. Să ne oprim puțin la definiția mediei aritmetice.

$$\overline{X} = \frac{1}{N}\sum X_i$$

Această mărime se mai numeşte şi moment de ordin 1, valorile individuale X_i fiind ridicate la puterea **1** şi apoi mediate.

Prin generalizare numim moment de ordin r marimea:

$$\overline{X} = \frac{1}{N}\sum X_i^r \qquad \text{(II.2.7.d)}$$

Pentru r = 2, $\overline{X}^2$ este media pătratică, pentru r = 3, $\overline{X}^3$ este media cubică, pentru r = -1, $\overline{X}_h$ este media armonică.

Dacă în locul valorilor individuale folosim abaterile centrale, momentele obținute se vor numi momente centrate, deci pentru momentul centrat de ordin **r** avem formula:

$$m_{cr} = \frac{1}{N}\sum (X_i - \overline{X})^r \qquad \text{(II.2.7.e)}$$

Observăm că pentru **r = 1** avem $m_{c1} = 0,$ iar pentru **r = 2** obținem $m_{c2} = s^2$ (dispersia).

Din cele relatate până aici putem sesiza că momentele de ordin 1 dau informații asupra indicatorilor tendinței centrale, iar cele de ordin 2, asupra indicatorilor de dispersie. Celelalte momente ne dau informații utile; să le analizăm pe scurt.

2.4. ASIMETRIA

Momentele de ordin 3 dau informații asupra simetriei distribuției.

a. Se defineşte un parametru numit "indice de asimetrie" (engl. skewness) prin relația:

$$m_{c3} = \frac{1}{N}\sum (x_i - \bar{x})^3 \qquad \text{(II.2.10.d)}$$

Pentru $m_{c3} = 0$ distribuția este simetrică,

$m_{c3} \langle 0$ asimetrie la stânga (II.2.10.b)

$m_{c3} \rangle 0$ asimetrie la dreapta (fig. II.2)

Pentru aprecierea asimetriei s-au propus şi alte relații:

b. Coeficientul de asimetrie Pearson:

$$\propto = \frac{M_0 - \overline{X}}{S} \qquad \text{(II.2.10.c)}$$

unde $\overline{X}$ = valoarea medie

M_0 = moda

s = deviația standard

După $\propto$ distribuția este simetrică sau asimetrică la stânga/dreapta la fel ca după m_{c3}. (fig. II.7.)

c. Coeficientul de asimetrie intercuantil:

$$\propto = \frac{(Q_3 - M_e) - (M_e - Q_1)}{(Q_3 - M_e) + (M_e - Q_1)} = \frac{Q_3 + Q_1 - 2M_e}{Q_3 - Q_1} \qquad \text{(II.2.10.d)}$$

având valori între +1 şi -1, cu 0 pentru distribuții simetrice.

d. Coeficientul bazat pe momentele centrate:

$$\propto = \frac{m_{c3}^2}{m_{c2}^2} \qquad \text{(II.2.10.e)}$$

cu interpretări similare cu indicele de asimetrie.

2.5. EXCESUL

Excesul este un parametru ce dă informații asupra gradului de turtire/boltire (limba engleza *kurtosis*). Se calculează cu relația:

$$\beta = \frac{m_{c4}}{m_{c2}^2} \qquad \text{(II.2.11.a)}$$

unde m_{c4} este momentul centrat de ordin 4 dat de:

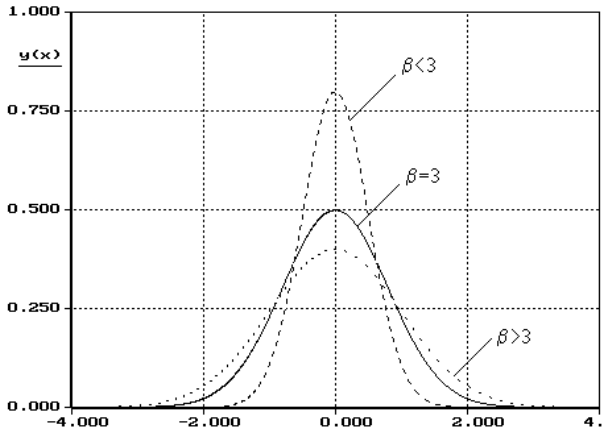


Figura II.5. Excesul (turtirea / boltirea)

$$m_{c4} = \frac{1}{N}\sum (X - \bar{X})^4 \qquad \text{(II.2.11.b)}$$

Pentru distribuția normală β = 3

Alte distribuții: distribuții mai turtite (β > 3) sau distribuții mai boltite (β < 3) - (fig. II.8).

3. DISTRIBUȚII

Am utilizat frecvent termenul de distribuție fără să ne ocupăm detaliat de el. Încercăm în cele ce urmează să aducem câteva precizări.

3.1. FUNCȚIA DE DISTRIBUȚIE

Definiție: Dacă x este o variabilă independentă, reprezentând valorile posibile ale unui parametru urmărit într-un studiu statistic atunci funcția

y = f (x), cu $y_i = p(x = x_i)$

care ne arată probabilitatea de a întâlni valoarea x într-o populație se numeşte **funcție de distribuție.**

Observatii:

- uzual se foloseşte nu o funcție continuă ci una discretă, în care valoarea funcției reprezintă probabilitatea de a întâlni mărimea studiată într-un interval (x_i , x_{i+1})
- vom face distincție între distribuțiile teoretice (în care calculăm valorile funcției) şi cele experimentale (în care valorile funcției au rezultat în urma unor măsurători).

3.2. FUNCȚII DE DISTRIBUȚIE UZUALE

În cele ce urmează vom aminti doar trei funcții de distribuție mai des întâlnite:

a. ***Distribuția uniformă***

$$f_{(xi)} = p(x = x_i) = k \qquad \text{(II.3.1.b)}$$

Experimental se obțin diverse fluctuații (fig. II.6.a)
Exemplu: probabilitatea de a arunca cu zarul valorile 1-6 este 1/6 pentru fiecare aruncare. După 100 de aruncări obținem o situație ca în fig. II.6.a.

b. ***Distribuția normală*** descrisă de formula (II.2.6.a) având forma unui clopot. Reprezentarea în fig. II.6.b.

c. ***Distribuția binomială*** *(utilă în studiul variabilelor calitative)*
Exemplu: probabilitatea de a extrage o bilă albă dintr-o urnă cu N bile dintre care A bile albe şi B bile negre (A+B=N) este: p = A / N. După extragere bila se introduce înapoi în urnă (Bernoulli).
Dacă din urnă se scot n bile atunci numărul x de bile albe extrase are o repartiție binomială:

$$f_{(x)} = C_n^x \cdot p^x \cdot q^{n-x} \qquad x = 0,1, ..., n \qquad \text{(II.3.1.c)}$$

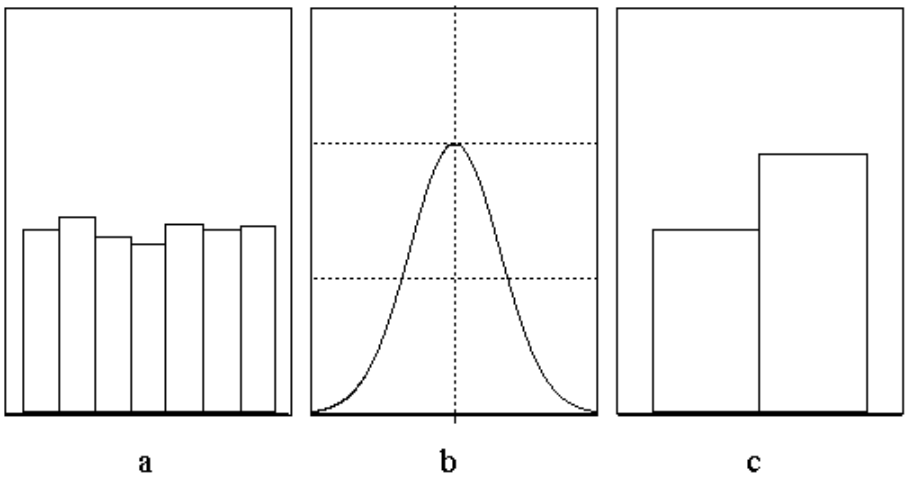


Figura II.6 Distribuții: a: uniformă, b: normală, c: binomial

Funcția de repartiție

Uneori, în loc de funcția de distribuție, care ne dă probabilitatea p_1 ca variabilă studiată x să aibă o anumită valoare x_1 (sau încadrată într-un interval în jurul lui x_i), se folosește o altă funcție numită *funcție de repartiție*:

$$y_r = p(x \le x_i) \qquad \text{(II.3.1.d)}$$

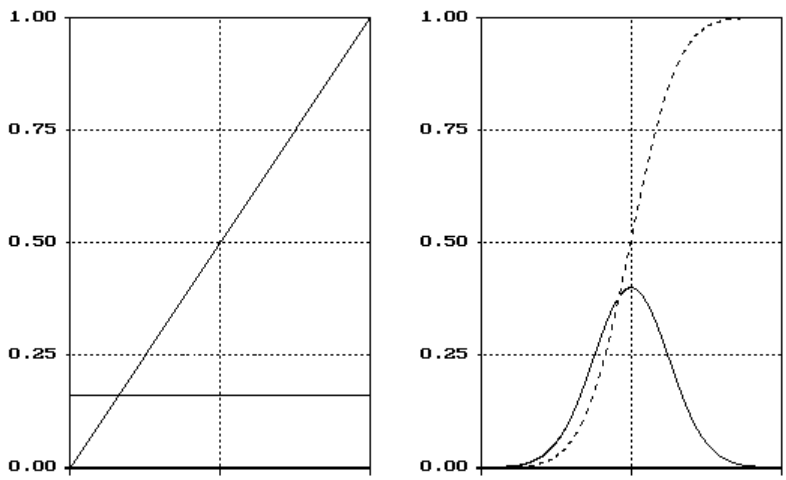


Figura II.7. Funcția de repartiție pentru distribuția uniformă (stanga) şi normală (dreapta) – funcțiile de repartiție sunt redate cu linie întreruptă, iar cele de distribuție cu linie continuă

În cazul funcțiilor experimentale discrete, funcția de repartiție ne dă frecvențele cumulate pentru toate clasele inferioare. Pentru distribuțiile uniformă şi normală, funcțiile de de repartiție sunt redate în figura II.7.

Distribuţia Gauss normalizată

Funcţia de distribuţie normală (Gauss) dată de formula (II.2.6.a) este simetrică faţă de μ. Dacă am face o schimbare de variabilă $y = x - \mu$ ea ar deveni simetrică faţă de origine şi ar mai depinde numai de un parametru: σ. Dacă am mai face încă o schimbare de variabilă, (practic alegând unităţi de măsură convenabile), putem obţine $\sigma = 1$, deci pentru

$$Z = \frac{x - \mu}{\sigma} \qquad \text{(II.3.1.e)}$$

obţinem curba de distribuţia Gauss normalizată (media=0, deviaţia standard = 1), sau distribuţia Z:

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} \qquad \text{(II.3.1.f)}$$

Deoarece funcţia Gauss nu se poate integra, valorile funcţiei de repartiţie pentru forma normalizată se găsesc în tabele în cărţile de statistică; de asemenea, programele de prelucrări statistice calculează aceste valori.

În afară de distribuţiile pomenite până acum există numeroase alte tipuri de distribuţii pe care le întâlnim în cercetare şi în practica biomedicală.

4. ESTIMAREA STATISTICĂ

Am văzut în paragrafele precedente că studiul unei populaţii se efectuează practic pe o submulţime a sa, pe un eşantion, iar concluziile obţinute pe un eşantion le vom extinde - prin operaţia de inferenţă statistică - la nivelul întregii populaţii. În timp ce concluziile noastre reprezintă afirmaţii adevărate efectiv doar pentru eşantion, la nivelul populaţiei ele au aceeaşi valoare de adevăr numai cu o anumită probabilitate - vom spune că "estimăm" parametrii populaţiei pornind de la valorile obţinute pe eşantion. Pentru a păstra rigurozitatea exprimărilor consacrate în acest domeniu, vom preciza câţiva termeni uzuali.

4.1. NOŢIUNEA DE ESTIMATOR

a. Terminologie

Mărimile caracteristice ale unei populaţii se numesc **parametri** şi au de obicei ca simboluri literele greceşti. Exemple: media populaţiei **μ,** deviaţia standard a populaţiei **σ ,** eroarea standard a mediei $\sigma_{\bar{x}}$**.**

Mărimile carcteristice ale unui eşantion se numesc **statistici** sau **indicatori** şi au de obicei simboluri litere latine. Exemple: media eşantionului $\overline{X}$, deviaţia standard **S**, eroarea standard a mediei $S_{\bar{x}}$.

Într-un studiu noi nu cunoaştem parametrii populaţiei ci doar determinăm statisticile eşantionului, fiecare folosită pentru a aproxima câte un parametru al populaţiei şi numită **estimator.** De exemplu: Vom spune că media eşantionului $\overline{X}$ este un estimator al mediei populaţiei **μ.** Deseori estimatorii se notează cu simbolul ^:

$\hat{x} = \bar{x} = est(\mu)$ sau $\hat{s} = s = est(\sigma)$.

Valoarea pe care o are un estimator într-o determinare concretă se numeşte **estimaţie.** De exemplu în studiul asupra dezvoltării copiilor, estimatorul pentru "înălţimea medie a copiilor de 10 ani " era " media înălţimii copiilor din eşantion" iar determinarea concretă avea estimaţia 137 cm.

b. Tipuri de estimări

Conform definiţiilor de mai sus un estimator aproximează în general valoarea unui parametru; în acest caz el se numeşte **estimator punctual.**

Caracteristica esenţială a unui estimator punctual este cea de a fi "nedeplasat".

Având în vedere faptul că prin inferenţă noi nu mai păstrăm o încredere deplină în estimările punctuale, pentru exprimarea probabilistă a încrederii rămase prin inferenţă vom încerca localizarea parametrului într-un interval; vom numi aceste aproximaţii **estimări prin intervale.** Aceste estimări sunt cele mai uzuale, iar în cele ce urmează ne vom referi numai la ele.

Lărgimea intervalului de încredere în care vrem să localizăm un parametru este dependentă de probabilitatea "încrederii" pe care o dorim şi anume: cu cât probabilitatea de a localiza parametrul este mai mare, cu atât intervalul de încredere este mai larg, deci probabilitatea de a greşi este mai mică. Însă creşterea nivelului de încredere nu ne foloseşte prea mult dacă intervalul devine atât de larg încât nu ne mai furnizează informaţii. Folosind exemplul discutat anterior observăm că dacă localizăm media populaţiei cu un nivel de încredere de 68%, intervalul este îngust: (136,138); pentru 95% devine (135,139), la 99,7% este (134,140) ş.a.m.d. Este deci nevoie de a găsi un compromis, un nivel de încredere care să ofere atât o localizare satisfăcătoare cât şi o probabilitate ridicată de a fi adevărată localizarea estimată, deci o probabilitate mică de a localiza greşit parametrul. Practica a demonstrat că un nivel de încredere de 95% satisface optim cerinţele în majoritatea cazurilor concrete. De aceea vom considera în continuare că nivelul de încredere este de 95% (exceptând cazurile în care vom menţiona în mod expres că am ales altă valoare). Atragem însă atenţia că acest nivel este **convenţional.**

Probabilitatea de a localiza greşit parametrul analizat se notează cu α şi se poate exprima nivelul de încredere (n.i.) prin relaţia:

$$n.i. = 1-\alpha \qquad \text{(II.4.1)}$$

c. Principalele tipuri de estimări

Deşi ne-am oprit până acum în exemplele discutate asupra estimării mediei populaţiei, în diferite studii ne putem concentra atenţia asupra unei palete mai largi de parametri. În cadrul cursului ne vom opri la estimarea următorilor parametri:

- media populaţiei - în cazul eşantioanelor mari / mici
- diferenţe între medii
- proporţia unei / unor clase
- diferenţe între proporţii.

4.2. ESTIMAREA MEDIEI POPULAŢIEI

A. Pentru eşantioane mari ($N \geq 30$)

a. Distribuţia mediilor eşantioanelor

În exemplul cu studiul dezvoltării copiilor, în care am făcut referiri la înălţimile copiilor dintr-un eşantion, am afirmat că, dacă reprezentăm grafic mediile eşantioanelor extrase din aceeaşi populaţie obţinem o distribuţie normală numai pentru eşantioane mari.

b. Formule

În acest caz, pentru α = 5% (n.i. = 95%) avem:

$$\hat{\mu} \in \left(\overline{X} - 1{,}96\, S_{\bar{x}} \; ; \; \overline{X} + 1{,}96\, S_{\bar{x}} \right) \qquad \text{(II.4.2.a)}$$

Valoarea 1,96 reprezintă valoarea funcției Z (distribuția Gauss normalizată) pentru a cuprinde în intervalul de mai sus 95% din arie (am văzut în cursul precedent că pentru Z = 2 cuprindeam 95,4% din aria de sub curbă) (fig. II.10.).

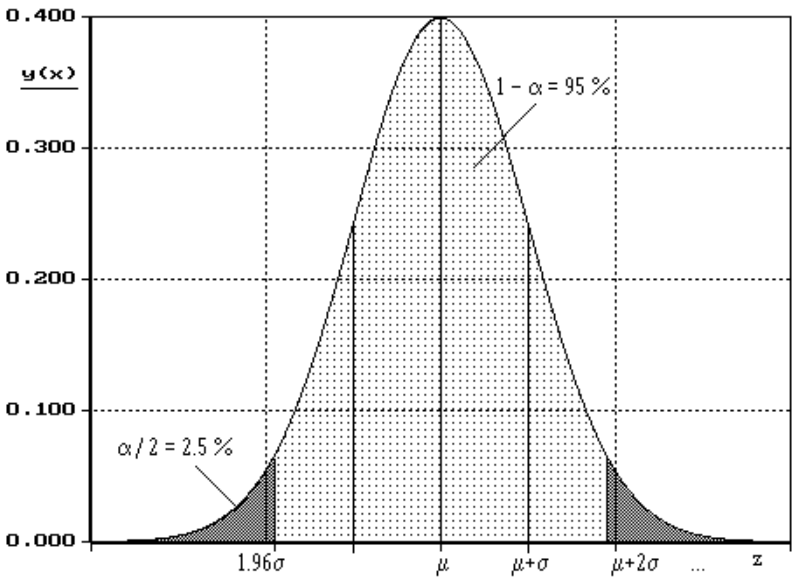


Figura II.10. Relația lui Z cu pragul de semnificație α şi cu nivelul de încredere 1 - α

Această valoare se mai notează $Z_{\alpha/2}$ indicele având aici semnificația că aria rămasă neinclusă este $\alpha/2$ = 2,5% (câte 2,5% în fiecare parte). [Obs: în unele cărți se notează $Z_{1-\alpha/2}$]. Deci într-o formă mai 52eneral putem scrie:

$$\hat{\mu} \in \left(\overline{X} - Z_{\alpha/2} \cdot S_{\bar{x}} \; ; \; \overline{X} + Z_{\alpha/2} \cdot S_{\bar{x}} \right) \qquad \text{(II.4.2.b)}$$

iar pentru $Z_{\alpha/2}$ vom lua o valoare din tabelul II.3.

Tabel II.3. Valorile scorului Z al distribuției normale

Nivel incredere **(1 - α)**	**0,68**	**0,90**	**0,95**	**0,98**	**0,99**
Prag de semnificație α	0,34	0,10	0,05	0,02	0,01
$z_{\alpha/2}$	1,00	1,65	1,96	2,33	2,58

De obicei manualele conțin anexe în care sunt prezentate diverse tabele.

c. Exemplul II.2.

Pe un eşantion de 144 sportivi se găseşte pentru VEMS (volumul expirator maxim în 1 secundă) valoarea medie $\overline{X}$ = 4,84 şi deviaţia standard S = 0,36. Să estimăm în ce interval găsim media populaţiei cu nivel de încredere de 98%.

Avem: $S_{\bar{x}} = S / \sqrt{N} = 0{,}36 / \sqrt{144} = 0{,}03\ l$

Pentru 1 - α = 98% găsim $Z_{\alpha/2}$ = 2,33, deci:

$$\hat{X} \in (4{,}84 - 0{,}03 \cdot 2{,}33;\ 4{,}84 + 0{,}03 \cdot 2{,}33),$$

$$\hat{X} \in (4{,}84 - 0{,}07;\ 4{,}84 + 0{,}07), \text{ adică } \hat{X} \in (4{,}77;\ 4{,}91).$$

Cu alte cuvinte, avem încredere de 98% că adevărata medie a VEMS pentru sportivi să fie între 4,77 şi 4,91, ceea ce înseamnă ca probabilitatea ca media VEMS la sportivi să fie în afara acestui interval este sub 2%.

B. Pentru eşantioane mici (N < 30)

a. Distribuţia mediilor eşantioanelor

După cum am mai specificat anterior, distribuţia mediilor eşantioanelor poate fi considerată distribuţie normală numai în cazul eşantioanelor mari. În cazul eşantioanelor mici (considerate convenţional mici dacă N < 30), mediile eşantioanelor au o distribuţie "t" (sau distribuţie normală, însă mai turtită - figura II.11); curba este cu atât mai turtită (deci mai diferită de curba Gauss) cu cât eşantionul este mai mic.

b. Formule

Curba de distribuţie "t" depinde deci de mărimea eşantionului care va fi caracterizată printr-un parametru, notat cu ν, numit "număr de grade de libertate" şi dat de relaţia:

$$\nu = N - 1 \qquad \text{(II.4.3)}$$

Lărgimea intervalului în care localizăm media populaţiei va fi dată de relaţia:

$$\hat{\mu} \in (\overline{X} - t_{\alpha/2,\nu} \cdot S_{\overline{X}};\ \overline{X} + t_{\alpha/2,\nu} \cdot S_{\overline{X}}) \qquad \text{(II.4.4)}$$

Relaţia (II.4.4.) este foarte asemănătoare cu (II.4.2.b)

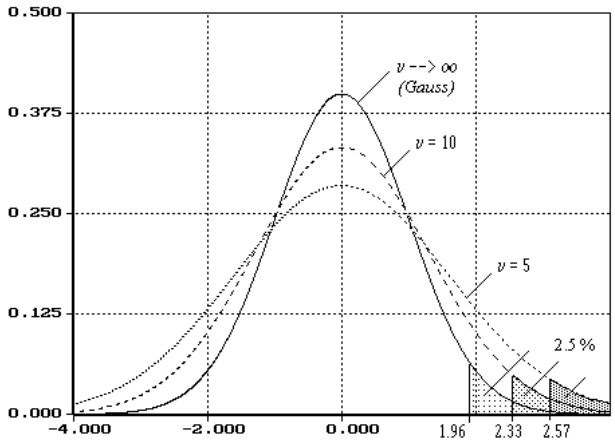


Figura II.11. Distribuţia "t". Valorile lui t care lasă câte 2,5% arie în fiecare parte (deci n.i. = 95%) sunt cu atât mai mari cu cât eşantionul este mai mic

Din tabelul II.4. se poate observa apropierea lui t de z pe măsură ce creşte dimensiunea eşantionului.

Tabel II.4. Valorile lui t pentru câteva nivele de încredere (1-α) şi grade de libertate (ν)

1-α ν	0.90	0.95	0.99
1	**6,31**	**12,7**	**63,7**
2	**2,92**	**4,30**	**9,92**
5	**2,01**	**2,57**	**4,03**
10	**1,81**	**2,23**	**3,17**
30	**1,70**	**2,04**	**2,75**
∝	**1,65**	**1,96**	**2,58**

c. Exemplul II.3

Considerăm din nou exemplul anterior, cu valoare medie a VEMS pe un lot de sportivi $\overline{X} = 4{,}84\,l$ şi $S = 0{,}36\ l$, dar să presupunem acum că am efectuat determinările pe un lot de numai N = 16 sportivi. Să vedem în ce măsură este influenţat intervalul în care putem localiza media populaţiei cu precizie de 98%. În primul rând: $S_{\overline{X}} = S/\sqrt{N} = 0{,}36/\sqrt{16} = 0{,}09$. Apoi pentru 1-∝ = 98% şi ν=16-1=15 grade de libertate găsim în tabelul distribuţiei **t** valoarea $t_{\alpha/2,\nu} = 2{,}60$ deci:

$$\hat{X} \in (4{,}84 - 0{,}09 * 2{,}60\,;\, 4{,}84 + 0{,}09 * 2{,}60)$$

$$\hat{X} \in (4{,}84 - 0{,}23\,;\, 4{,}84 + 0{,}23)$$

$$\hat{X} \in (4{,}60\,;\, 5{,}08).$$

Observăm că pierderea de precizie în localizare este imensă, intervalul fiind de peste 3 ori mai larg comparativ cu localizarea obţinută pe un eşantion mare.

De aceea, după cum vom vedea chiar în încheierea acestui subiect, în biostatistică putem calcula o dimensiune minimă a eşantionilui pentru a putea obţine localizări ale mărimilor estimate în intervale suficient de înguste şi cu nivel de încredere satisfăcător de ridicat.

4.3. ESTIMAREA PROCENTELOR

a. Distribuţia procentului în eşantion

În cazul variabilelor calitative (nominale), indivizii dintr-un eşantion sunt grupaţi în clase; distribuţia de acest gen se numeşte distribuţie binominală. În cazul unei populaţii avem deci câte un procent real π_i pentru fiecare clasă i:

1	2	...	i	...	k
π_1	π_2	...	π_i	...	π_k

(II.4.5)

La extragerea unui eşantion din populaţie vom obţine procentele $p_1, p_2, ..., p_k$, cu deviaţiile standard ale procentelor date de relaţia:

$$S_p = \sqrt{\frac{p(1-p)}{N}} \qquad \text{(II.4.6)}$$

Dacă repetăm extragerea eşantionului, fiecare procent p_i va prezenta variaţii. Pentru eşantioane mari procentele prezintă o repartiţie normală (pentru procente nici nu se poate lucra pe eşantioane mici!).

b. Formule

Vom putea deci aplica "scorul Z", la fel ca în cazul mediilor, deci:

$$\hat{p}_i \in \left(p_i - Z_{\alpha/2} \cdot S_{pi} \; ; \; p_i + Z_{\alpha/2} \cdot S_{pi} \right) \qquad \text{(II.4.7)}$$

c. Exemplul II.3

Dintr-un lot de 80 de indivizi, 24 au grupa sanguină A. Care este proporţia reală a grupei sanguine A în populaţia studiată cu nivel de încredere de 95%.

$$p = \frac{100 \cdot 24}{80} = 30\% \; ; \; s_p = \sqrt{\frac{0.3 * 0.7}{80}} = 0.0512 \equiv 5.12\%$$

$\hat{p} \in$ (30 – 1.96*5.12; 30 + 1.96*5.12) sau

$\hat{p} \in$ (30 - 10, 30 + 10), adică $\hat{p} \in$ (20% ; 40%)

Cu alte cuvinte, din studiul efectuat putem face doar afirmaţia că procentul de răspândire al grupei A este între 20% şi 40%, cu nivel de încredere de 95%.

Dacă aceeaşi proporţie, de 30% o găseam pe un lot de 800 de indivizi (240 din 800), obţineam $S_p = 1.62\%$ şi $\hat{p} \in$ (26,8%; 33,2%).

4.4. ESTIMAREA DIFERENŢELOR

În numeroase studii urmărim nu atât valorile absolute ale unor parametri, care au împrăştieri naturale destul de largi, ci în special variaţiile mărimilor. Aceste variaţii pot fi urmărite atât pentru valorile propri-zise (medii ale eşantioanelor), cât şi pentru proporţiile din eşantioane ce aparţin unei clase.

A. Diferenţe între medii

a. Pentru loturi diferite

Exemplu II.4. care este diferenţa de înălţime între băieţii şi fetele de 10 ani?

Evident, răspunsul se dă după un studiu în care obţinem, pe două loturi, valori de genul:

$$n_B = 25, \overline{X}_B = 137{,}2\,cm, S_B = 5{,}1\,cm$$

$$n_F = 25, \overline{X}_F = 138{,}6\,cm, S_F = 5{,}1\,cm$$

Estimarea diferenţei se face prin:

$$d_{\overline{X}_{B-}} = \overline{X}_B - \overline{X}_F = -1{,}4\,cm \qquad \text{(II.4.8.)}$$

Intervalul de încredere se apreciază cu ajutorul deviaţiei standard estimate prin diferenţe:

$$S_d = \sqrt{\frac{S_B^2}{n_B} + \frac{S_F^2}{n_F}} \qquad \text{(II.4.9.)}$$

pentru loturi mari ($n_{1,2} \rangle 30$) încadrarea o vom face după

$$\hat{d}_{\bar{x}-\bar{x}_2} \in (\bar{d} - Z_{\alpha/2} \cdot S_d ;\ \bar{d} + Z_{\alpha/2} \cdot S_d) \qquad \text{(II.4.10.a)}$$

pentru loturi mici ($n_{1,2} < 30$)

$$\hat{d}_{\bar{x}_1-\bar{x}_2} \in (\bar{d} - t_{\alpha/2,\nu} \cdot S_d ;\ \bar{d} + t_{\alpha/2,\nu} \cdot S_d) \qquad \text{(II.4.10.b)}$$

b. Pentru serii perechi

Vom considera în continuare un caz aparte, întâlnit destul de des. Să începem cu un exemplu.

Exemplul II.5. Un medicament antihipertensiv poate fi testat fie considerând două loturi - unul tratat şi unul martor - fie lucrând pe un singur lot şi facând un set de măsurători ale tensiunii arteriale înainte de tratament, respectiv după tratament. Este perferată ultima variantă. În acest caz vom avea, pentru fiecare individ **i** din lot efectul exprimat sub formă de diferență dintre două valori:

$$d_i = X_{2i} - X_{1i} \qquad \text{(II.4.11.)}$$

unde X_{2i} este valoarea după tratament iar X_{1i} este valoarea înainte de tratament.

Eroarea standard a diferențelor este dată de relația:

$$S_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n(n-1)}} \qquad \text{(II.4.12.)}$$

Estimarea diferenței pentru serii perechi va fi dată tot de relațiile (II.4.10.a) sau (II.4.10.b).

B. Diferențe între procente

Asemănător cu raționamentele deja prezentate până acum, putem încadra şi estimarea diferenței a două procente:

$$d_p = p_2 - p_1 \qquad \text{(II.4.13.)}$$

Pentru eroarea standard a diferenței a două procente folosim formula:

$$s_{pd} = \sqrt{\frac{p_1(1-p_1)}{n_1 - 1} + \frac{p_2(1-p_2)}{n_2 - 1}} \qquad \text{(II.4.14.)}$$

Intervalul de încredere al estimării va fi:

$$\hat{d}_p \in (\bar{d} - Z_{\alpha/2} \cdot S_{pd} ;\ \bar{d} + Z_{\alpha/2} \cdot S_{pd}) \qquad \text{(II.4.15.)}$$

4.5. CALCULUL MĂRIMII EŞANTIONULUI

Cea mai importantă consecință a studiilor privind încadrarea unui parametru într-un interval este calculul mărimii eşantionului. Am văzut că lărgimea intervalului în care încadrăm estimația depinde puternic de dimensiunea eşantionului **(n)**. Această dependență ne poate folosi la evaluarea mărimii eşantionului astfel încât să obținem încadrarea parametrului populației într-un interval rezonabil de îngust.

Exemplul II.6. Dorim să determinăm înălțimea medie a copiilor de 10 ani cu precizie de ± 1 cm , având un nivel de încredere de cel puțin 95%. Ce dimensiune minimă trebuie să aibă eşantionul?

Conform relației (II.4.2.b) vom avea:

$Z_{\alpha/2} \cdot S_{\bar{x}} = 1{,}96 \cdot S_{\bar{x}} = 1$ *cm,* adică*:* $S_{\bar{x}} = 1/1{,}96 \cong 0{,}5$ *cm.*

Pentru calculul lui **n** ar trebui cunoscut gradul de împrăștiere a valorilor înălțimii pentru populație, exprimat prin deviația standard a populației σ ; de obicei această mărime nu este cunoscută şi în locul ei se foloseşte o estimare a deviației standard **S** obținută într-un studiu pe un eşantion. În cazul exemplului nostru, considerând că într-un studiu anterior s-a găsit S = 6 cm, din relația (5.8.b) obținem:

$$n = (S / S_{\bar{x}})^2 \qquad \text{(II.4.16)}$$

în care înlocuind valorile din exemplu

$n = (6 / 0{,}5)^2 = 144.$

Deseori nici **s** nu este cunoscut şi atunci fie facem un studiu preliminar fie îl aproximăm pe **s** cu relația:

$$S = (X_{max} - X_{min})/6 \qquad \text{(II.4.17)}$$

în care X_{min} şi X_{max} reprezintă valorile extreme la care ne-am aştepta, evaluate conform experiențelor noastre anterioare.

5. TESTE STATISTICE

5.1. NOȚIUNI GENERALE

A. Diferențe semnificative şi nesemnificative din punct de vedere statistic

În exemplul nostru cu înălțimea copiilor nu am făcut până acum distincție între băieți şi fete. Un studiu asupra dezvoltarii copiilor ar trebui să țină cont de evoluția hormonală diferită care va genera dezvoltarea somatică diferită. Vom dezvolta exemplul nostru astfel:

Exemplul II.7. Pe un lot de 36 de băieți de 10 ani obținem pentru înălțime următoarele rezultate: $n_B = 36$, $x_B = 137cm$, $S_B = 12$ *cm,* deci $S_{\bar{x}}^B = 2cm$, iar pe un grup de 36 fete de aceeaşi vârstă: $n_F = 36$, $\overline{X}_F = 140cm$, $S_F = 12$ *cm, deci* $S_{\bar{X}}^F = 2$ *cm.*

Întrebare: diferențele observate arată că fetele de 10 ani sunt mai înalte decât băieții sau pot să fie atribuite întâmplării?

Să remarcăm mai întâi că diferențele observate în cursul unor studii pot fi clasificate în două categorii:

1. - diferențe ce pot fi atribuite întâmplării (variabilității de eşantionare); acestea vor fi numite **diferențe nesemnificative**

2. - diferențe ce pot avea alte cauze numite **diferențe semnificative.**

Să analizăm datele de mai sus: pentru un nivel de încredere de *95%, $Z_{0,95}$ = 2,02* ≈ 2 deci $\hat{\mu}_B \in (133,141)$, cu alte cuvinte avem probabilitatea de 95% ca media unui alt eşantion, extras din aceeaşi populație (de băieți) să se găsească în intervalul *133 - 141* cm. Faptul că pentru un eşantion, cum ar fi grupul de fete, am găsit media *140* s-ar putea datora deci**, în mare măsură,** întâmplării. Accentuăm exprimarea "în mare măsură", fiindcă aici apare un arbitrar; noi am apreciat intervalul *(133; 141)* considerând un nivel de încredere convențional de 95%. Acceptând această convenție ajungem la concluzia că, dacă media înălțimii fetelor este de 140, acest lucru nu ne îndreptățeşte să afirmăm că fetele sunt mai înalte fiindcă sunt şanse mari ca diferențe de acest gen să apară din întâmplare. În schimb, dacă obținem pentru fete înălțimea medie $\overline{X}_F^1 = 142$ cm, această valoare cădea în afara intervalului *(133 ; 141);* probabilitatea ca

să obţinem din întâmplare media unui eşantion în afara acestui interval este sub 5%; în acest caz noi vom considera că nu din întâmplare s-a obţinut această valoare şi că diferenţele sunt semnificative (fig. II.12).

Să facem două observaţii importante:

1. - pragul de 5% pe care l-am folosit în exemplul nostru pentru a decide dacă diferenţele vor fi considerate semnificative sau nesemnificative este convenţional; el se numeşte **prag de semnificaţie** şi este considerat 5% în majoritatea cazurilor;

2. - concluziile au un caracter pur probabilist; dacă obţinem $\overline{X}_F^I$ şi spunem că fetele sunt mai înalte, să nu uităm că există o anumită probabilitate, chiar dacă este mică (sub 5%) ca să fi obţinut asemenea valori din întâmplare, adică în realitate să nu avem diferenţe semnificative.

De asemenea faptul că pentru cazul $\overline{X}_F = 140$ spunem că diferenţele de înălţime între băieţi şi fete nu sunt semnificative din punct de vedere statistic nu însemnă că **în realitate** nu avem diferenţe semnificative - numai faptul că din măsurătorile efectuate nu putem afirma că ar fi "statistic semnificative".

De aceea în analiza statistică pe care o efectuăm când aplicăm testele statistice pornim cu formularea unei ipoteze, pe care o vom accepta sau respinge cu o anumită probabilitate.

5.2. IPOTEZE STATISTICE

Testele statistice prin care se realizează o comparaţie încep cu enunţarea unei ipoteze privind un rezultat posibil al comparţiei, numită ipoteză statistică, pe care o putem defini astfel:

Definiţie*:* Ipoteza statistică este o propoziţie conţinând o afirmaţie sau negaţie privind un parametru al unei populaţii sau o lege de distribuţie.

Ipotezele au diferite variante de formulări, de aceea le vom defini la început în cazuri nu foarte generale.

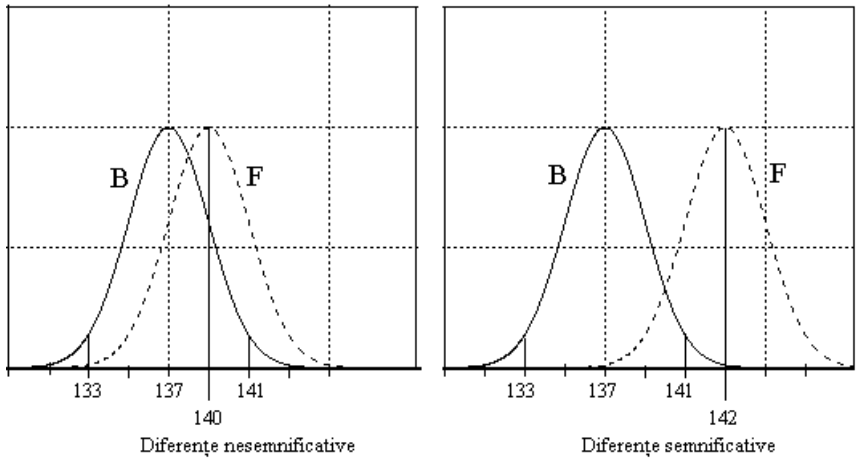


Figura II.12. Ilustrarea diferenţelor nesemnificative şi semnificative din exemplul II.7

A. Ipoteza de zero

Ipoteza de zero face întodeauna afirmaţia ca "între elementele pe care le comparăm nu există diferenţe semnificative".

Ipoteza de zero se notează prescurtat de obicei cu H_0 şi se mai numeşte **ipoteză de nul sau ipoteză nulă** (impropriu).

Exemplu: În cazul în care comparăm înălțimea medie a băieților şi fetelor, ipoteza de zero s-ar scrie:

$$H_0: \overline{X}_B = \overline{X}_F \qquad \text{(II.5.1)}$$

Alte enunțuri echivalente:
- "diferențele observate se datoresc numai întâmplării"
- "nu putem afirma că între cele două valori (serii, distribuții) există diferențe semnificative".

Prin ipoteza de zero putem compara:
- o valoare medie (obținută pe un lot) cu valoare dată (număr)
- două valori medii (două loturi) - cel mai adesea
- o distribuție experimentală cu una teoretică
- două distribuții experimentale
- două dispersii
- mai multe valori medii
- mai multe dispersii, etc., etc.

B. Ipoteze alternative

Propozițiile care sunt adevărate când/dacă H_0 nu este adevărată se numesc ipoteze alternative şi se notează cu H_a *sau* H_1.

Ipotezele alternative se pot cel mai simplu exemplifica în cazul comparării a două valori medii $\overline{X}_B$; $\overline{X}_F$. Vom putea avea situațiile:

a. $\overline{X}_B \neq \overline{X}_F$ ("înălțimea băieților este diferită de a fetelor")

b. $X_B \rangle \overline{X}_F$ ("băieții sunt mai înalți decât fetele")

c. $\overline{X}_B \langle \overline{X}_F$ ("fetele sunt mai înalte decât băieții).

Ipoteza alternativă (a) se numeşte **bilaterală** (este adevărată atât când $\overline{X}_B \rangle \overline{X}_F$ cât şi când $\overline{X}_B \langle \overline{X}_F$), în timp ce variantele (b) şi (c) se numesc **unilaterale.**

C. Prag de semnificație

Am văzut că pentru a putea stabili dacă diferențele sunt semnificative sau nu, trebuie să ne alegem arbitrar un prag al probabilității, numit **prag de semnificație, α,** cu ajutorul căruia stabilim lățimea intervalului în care considerăm că avem fluctuațiile, datorate întâmplării (atribuite variabilității de eşantionare); dacă valoarea de comparat va fi inclusă în acest interval vom spune că diferențele sunt nesemnificative şi vom accepta ipoteza de zero. De aceea acest interval se mai numeşte **regiune de acceptare;** limitele intervalului se numesc **valori critice**, iar regiunea exterioară se numeşte **regiune de respingere** sau **regiune critică** (fig. II.13).

Pragul de semnificație α are o valoare arbitrară. Alegerea lui trebuie să satisfacă două condiții:

- pe de o parte valoarea trebuie să fie suficient de mică pentru ca probabilitatea ca din întâmplare să obținem diferențe la fel de mari să fie redusă

- pe de altă parte alegerea unei valori prea mici lărgeşte prea mult regiunea de acceptare, deci α trebuie să fie suficient de mare pentru a menține intervalul destul de îngust.

Convenţional se acceptă că, în majoritatea studiilor în medicină şi biologie, este satisfăcătoare valoarea $\alpha = 0,05$.

D. Nivel de încredere

Valoarea 1 - α se numeşte nivel de încredere şi se exprimă de obicei în procente. Deci pentru valoarea uzuală $\alpha = 0.05$ nivelul de încredere este de 95%. Cu alte cuvinte, în acest caz avem o încredere de 95% (probabilitate de 95%) ca decizia pe care o luăm prin aplicarea testului să fie corectă.

5.3. ETAPELE APLICĂRII TESTULUI STATISTIC

După ce am definit principalele mărimi folosite pentru aplicarea unui test statistic putem sistematiza etapele de lucru:

A. Definirea mărimilor de comparat - evident, trebuie precizat la începutul studiului care vor fi mărimile asupra cărora se îndreaptă atenţia şi asupra cărora se vor aplica testele. În funcţie de acestea vom alege diferite tipuri de teste.

B. Formularea ipotezei zero şi a celei alternative - operaţie primară, fiindcă rezultatul testului (decizia) se exprimă în funcţie de H_0: se acceptă sau se respinge. Dacă nu se urmăreşte în mod special o ipoteză alternativă H_{1B} sau H_{1C}, se acceptă ca ipoteză alternativă cea bilaterală H_{1A}.

C. Alegerea pragului de semnificaţie al testului - în majoritatea cazurilor se ia $\alpha = 0,05$ care conferă un nivel de încredere de 95%.

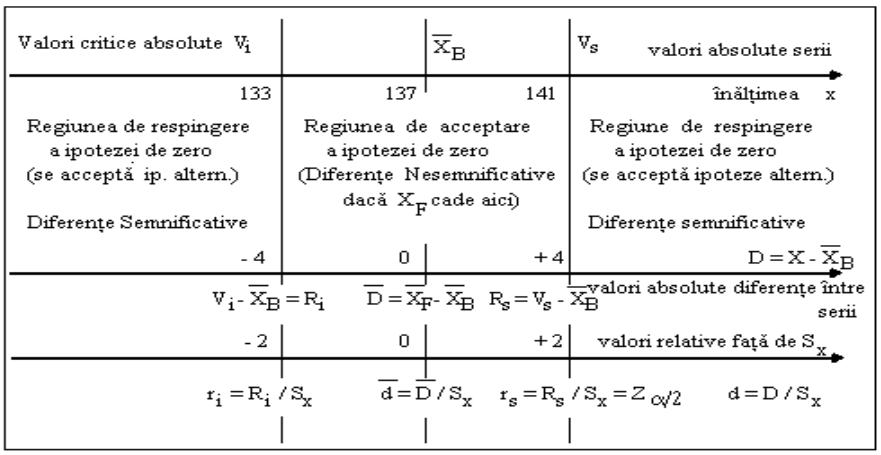


Figura II.13. Regiunea de acceptare şi H_0 exprimată cu valori absolute ale seriilor şi diferenţelor şi cu valori relative, pe scara normalizată faţă de $S_{\bar{X}}$. În aceste exemple s-a utilizat pragul de semnificaţie $\alpha = 0,05$.

D. Alegerea testului - este etapa esenţială căreia îi vom dedica un paragraf separat; în funcţie de tipul de variabile şi modul de distribuţie al valorilor se alege testul cel mai potrivit în funcţie de care se efectuează calculele (etapele E şi F).

E. Calculul valorilor critice (de obicei cele relative) şi stabilirea regiunilor de acceptare / respingere a ipotezei zero.

F. Calculul coeficientului **p** care reprezintă probabilitatea ca:
- ipoteza de zero să fie adevărată, sau
- diferenţele sa fie nesemnificative, sau
- să ne încadrăm în regiunea de acceptare.

Etapele **E** si **F** nu sunt distincte din punct de vedere al calculelor.

G. Formularea deciziei - etapă finală, în funcție de **p**:

- dacă $p \geq \alpha$ **acceptăm** H_0 şi spunem că diferențele sunt nesemnificative
- dacă $p < \alpha$ **respingem** H_0 şi spunem că diferențele sunt semnificative.

Pentru $\alpha = 0{,}05$ regiunea de respingere sew împarte la rândul ei în 3 subregiuni în care se încearcă o gradare a diferențelor semnificative (fig. II.14).

În toate cazurile în care diferențele sunt semnificative ipoteza de zero H_0 se respinge.

Observație: Statistic semnificativ nu înseamnă şi important din punct de vedere bio-medical.

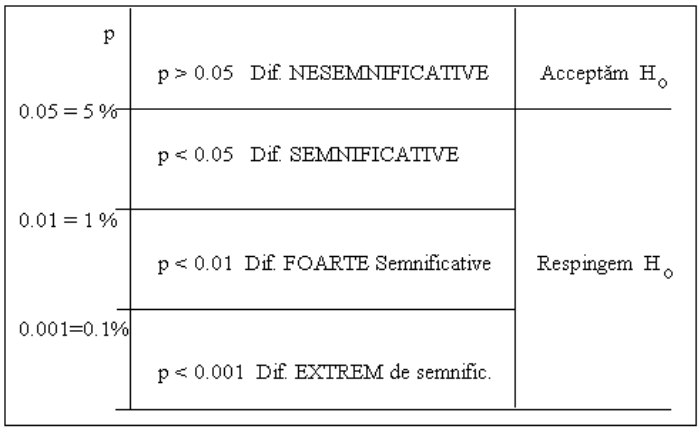


Figura II.14. Formularea deciziei unui test statistic în funcție de valoarea lui p

5.4. ERORI

Este foarte important sa nu pierdem din vedere faptul că decizia unui test statistic are caracter probabilistic. Deci faptul că în cazul $\overline{X}_B = 137cm$ *şi* $\overline{X}_F = 140cm$ am obținut $p>0{,}05$ şi am acceptat H_0 afirmând că diferențele sunt nesemnificative **nu înseamnă că în realitate** nu avem diferențe de înălțime între băieți şi fete la 10 ani ci doar faptul că, din studiul efectuat de noi, probabilitatea ca fetele şi băieții să aibă aceeaşi înălțime este mai mare decât 5%, ceea ce nu ne permite să afirmăm că diferențele sunt semnificative şi deci le-am putea atribui întâmplării (variabilității de eşantionare). S-ar putea ca în realitate diferențele să fie semnificative dar din diverse motive (fie pură întâmplare, fie loturi prea mici - din care cauză se obține o valoare mare pentru $S_{\overline{X}}$) aceste diferențe n-au fost sesizate ca atare. Există deci riscul de a avea erori în decizia noastră.

Erorile statistice posibile se împart în două clase:

a. erori de tip I : când respingeam H_0 deşi este **adevărată**
b. erori de tip II: când acceptăm H_0 deşi este **falsă.**

Probabilitatea erorii de tip I se notează cu α (este de fapt chiar legată de pragul de semnificație), iar cea a erorii de tip II cu β.

Situațiile posibile de decizie sunt sintetizate în tabelul II.5.1.

Tabelul II.5. Situații posibile în decizia testelor statistice

		SITUAȚIA	REALĂ
		H_0 = Adevărată	**H_0 = Falsă**
DECIZIE	Acceptăm H_0	**Corect** $p=1-\alpha$	**Eroare tip II** $p=\beta$
	Respingem H_0	**Eroare tip I** $p = \alpha$	**Corect** $p = 1 - \beta$

Observație: Denumirea de "erori de tip I şi II" este din ce în ce mai des întocmită cu cea de "risc de (eroare de) tip I sau II"; valorile α şi β arată doar "probabilitatea de a le comite".

5.5. CARACTERISTICILE TESTELOR STATISTICE

A. Nivelul de încredere

Definiție: Mărimea $1-\alpha$ se numeşte "nivel de încredere" sau "nivel de confidență" (uneori simplu "încrederea" sau "confidența") testului; α reprezintă pragul de semnificație, sau probabilitatea erorii de tip I şi reprezintă capacitatea de a accepta o ipoteză când aceasta este adevărată.

B. Puterea testului

Definiție: Mărimea $1 - \beta$ se numeşte "puterea testului", unde β reprezintă probabilitatea erorii de tip II şi reprezintă capacitatea de a respinge o ipoteză când aceasta este falsă.

Cele două caracteristici, nivelul de încredere şi puterea testului se află în relație de inversă proporționalitate. Într-adevăr, dacă am dori să creştem nivelul de încredere al testului, $1 - \alpha$, ar trebui micşorat pragul α, de exemplu de la 5% la 1%, în acest caz intervalul în care vom încadra media unui eşantion extras din aceeaşi populație va fi mai larg ($Z_{0,99} \approx 2,33$ față de $Z_{0,95} \approx 1,96$) deci suntem mai încrezători că, dacă este adevărat că cele două medii nu sunt semnificativ diferite (chiar dacă din întâmplare a apărut o diferență puțin mai mare), acesta nu va afecta decizia. Deci scade probabilitatea erorii de tip I.

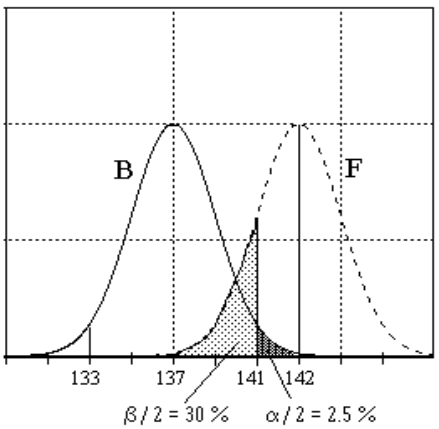


Figura II.15. Ilustrarea caracteristicilor unui test. În acest exemplu, valorile medii pentru populații sunt $\mu_B = 137$ *cm*, $\mu_F = 142$ *cm.* Luând referință lotul de băieți, pentru $\alpha = 5\%$ intevalul de acceptare este (133,141). Față de limita 141 avem probabilitatea de cca 30% de a găsi $\overline{X}_F < 141$ dar şi $p \approx 2,5\%$ pentru a găsi $\overline{X}_B > 141$.

În schimb, dacă în realitate diferenţele sunt statistic semnificative, această lărgire a intervalului ne poate conduce la concluzia falsă că diferenţele sunt nesemnificative, în timp ce ele în realitate sunt! Deci creşte probabilitatea erorii de tip II.

Relaţia între nivelul de încredere şi puterea testului poate fi ilustrată ca în figura II.15; construită pe baza datelor din exemplul folosit anterior.

5.6. TESTE PARAMETRICE ŞI NEPARAMETRICE

Stabilirea regiunii de acceptare este evident dependentă de tipul de distribuţie a valorilor (în exemplele discutate până acum s-a considerat ca îndeplinită distribuţia normală). Însă există situaţii în care nu cunoaştem tipul de distribuţie, astfel încât nu mai putem calcula atât de simplu regiunea de acceptare. În funcţie de acest aspect putem împărţi testele în două categorii:

a. *teste parametrice* - în care distribuţia este cunoscută (cel mai adesea se consideră doar distribuţia normală, pentru care se pot aplica aceste teste),

b. teste neparametrice - în care se consideră necunoscută distribuţia; testele neparametrice sunt mai generale; dacă distribuţia este în realitate o distribuţie normală testele neparametrice dau - în majoritatea cazurilor - rezultate asemănătoare cu cele parametrice; de aceea, în ultimul timp, ele se folosesc din ce în ce mai mult.

5.7. CLASIFICAREA TESTELOR STATISTICE

În funcţie de mărimile comparate putem distinge mai multe clase de teste statistice:

A. Teste de semnificaţie - prin care se verifică egalitatea unui parametru estimat (medie, procent, etc.) cu o valoare dată.

B. Teste de omogenitate - prin care se compară doi parametri (medii, procente, dispersii etc.).

Observaţie: în unele manuale sunt considerate teste de omogenitate numai cele de comparaţie a parametrilor de dispersie (deviaţii standard etc.), iar pentru compararea a două medii sau proporţii se utilizează termenul de "teste de semnificaţie".

C. Teste de concordanţă - prin care se compară o distribuţie experimentală cu una teoretică sau se compară două distribuţii experimentale.

D. Teste de independenţă - prin care se verifică independenţa unor serii de valori experimentale (în special pentru tabele de contingenţă).

E. Teste pentru corelaţii - prin care se evaluează semnificaţia parametrilor estimaţi în analiza corelaţiei.

Observaţie: Unii autori includ aceste teste în categoria testelor de semnificaţie.

Din punct de vedere teoretic se pot compara statistic şi alte elemente, specifice unui anumit domeniu (ex.: în analiza semnalelor biologice) pe care le vom prezenta în contextul corespunzător.

5.8. TESTE UZUALE ÎN BIOSTATISTICĂ

În paragraful care urmează vom descrie cele mai importante teste folosite în biostatistică pe care le vom prezenta pornind de la mărimile care se compară.

A. Se compară o **valoare medie** cu o **valoare dată**

. Ipoteza de zero**:** $\overline{X} = X_0$

. Test aplicat:

a. **testul Z -** dacă $n > 30$

b. **testul t** - dacă $n \leq 30$

Observaţie: Denumirea de "testul Z" nu este folosită prea des deoarece distribuţia normală Z este un caz limită al distribuţiei t, când numărul gradelor de libertate este foarte mare. Denumirea uzuală pentru testul aplicat în aceste condiţii va fi "testul t pentru o serie"

Observaţie: dacă σ_m (eroarea standard a populaţiei) este cunoscută, se foloseşte ea în calculul intervalului de acceptare (vezi formula 4.2.b. sau 4.4); dacă nu este cunoscută, se foloseşte estimatorul ei, $S_{\bar{X}}$ (formula 5.8.b).

Exemplu II.8 Într-un raport se susţine că înălţimea medie a copiilor de 10 ani este 139 cm. Acceptăm această afirmaţie?

Rezolvare: luăm un lot având $n = 36$ copii pe care obţinem $\bar{X} = 137{,}3$ *cm* , $S = 9$ *cm.*

. Ipoteza de zero: $H_0 : 137{,}3 = 139$ *cm*

. Alegem pragul de semnificaţie $\alpha = 5\%$; *atunci* $Z_\alpha = 1{,}96 \approx 2$

. Pentru lotul nostru $S_{\bar{X}} = 9/\sqrt{36} = 1{,}5$ *cm*

. Intervalul de acceptare este:

- în valori absolute : *(137,3 - 1,96 . 1,5; 137,3 + 1,96 . 1,5) adică (137,3 - 3; 137,3 + 3), sau (134,3 - 140,3);* valoarea $X_0 = 139$ se găseşte în acest interval, deci acceptăm H_0 şi spunem că diferenţele observate (între media experimentală 137,3 şi valoarea ipotetică 139) sunt nesemnificative şi se datoresc întâmplării.

- în valori absolute ale diferenţelor:

$\bar{D} = X_0 - \bar{X} = 139\text{-}137{,}3 = 1{,}7$ *cm*

intervalul fiind *(-1,96*1,5 ; +1,96*1,5)* adică *(-3, +3)*

Valoarea *1,7* este în acest interval, deci **acceptăm** H_0.

- în valori relative:

$\bar{d} = \bar{D} / S_{\bar{X}} = 1{,}7 / 1{,}5 \approx 1{,}13$

intervalul fiind *(-1,96; +1,96)*

*V*aloarea *1,13* fiind în acest interval, **acceptăm** H_0 **.**

B. Se compară **două valori medii**

. Ipoteza de zero: $H_0 : \bar{X}_1 = \bar{X}_2$

. Condiţii: $S_1 \approx S_2$; se poate $N_1 \neq N_2$

. Grade de libertate: $\nu = N_1 + N_2 - 2$

. Test aplicat:

a) Parametric: testul **t nepereche** (testul Student)

b) Neparametric: testul **Mann - Whitney**

În continuare prezentăm două variante de raţionament:

Varianta I

Exemplu II.9. Se analizează capacitatea vitală a unui grup de sportivi comparativ cu un grup de control, obţinând:

. lot sportivi: $N_1 = 36$, $\bar{X}_1 = 5{,}39$ *l* , $S_1 = 0{,}60$ *l*

. lot martor: $N_2 = 50$, $\bar{X}_2 = 4{,}83$ *l*, $S_2 = 0{,}70$ *l*

Ipoteza de zero: $H_0 : 5{,}39\ (\pm 0{,}6) = 4{,}83(\pm 0{,}7)$

Alegem pragul de semnificaţie: $\alpha = 5\%$.

Numărul gradelor de libertate: $\nu = 36 + 50 - 2 = 84$

Din tabelul cu valorile distribuției t, observam ca pentru valori între 60 şi 120 grade de libertate, coeficientul de încredere va fi acelaşi. Deci, pentru un test bilateral (*two-tailed*), găsim $t_{0,975;60} = 2.00$. Fiind o valoare din tabel, o vom nota în continuare cu indicele *"t"* deci $t_t = 2.00$.

Calculăm eroarea standard pentru diferențe:

$$S_d = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \qquad \text{(II.5.2)}$$

$$S_d = \sqrt{\frac{0.36}{36} + \frac{0.49}{50}} = \sqrt{0.02} = 0{,}14$$

Deci intervalul de acceptare a diferențelor este:

$\bar{d} \in (-t_t \cdot S_d, +t_t \cdot S_d)$, adică: (II.5.3.)

$\bar{d} \in \langle -2*0.14; +2*0.14 \rangle$ sau $\bar{d} \in \langle -0.28; 0.28 \rangle$

Diferența obținută este:

$$\bar{d} = \bar{X}_1 - \bar{X}_2 = 0.56 \qquad \text{(II.5.4.)}$$

deci este situată în afara regiunii de acceptare şi noi vom **respinge H_0**, considerând adevărata ipoteză alternativă şi vom spune că diferențele sunt semnificative.

Pentru a vedea eventual şi "cât de semnificative" sunt aceste diferențe, putem calcula regiunile de acceptare pentru:

. foarte semnificative: $t_t = t_{0,99:60} = 2{,}66;\ \bar{d} \in (-0{,}37; +0{,}37)$

. extrem de semnificative: $t_t = t_{0,999:60} = 3{,}37;\ \bar{d} \in (-0{,}47; +0{,}47)$

Observăm că diferența reală $\bar{d} = 0{,}56$ este în afara atât a intervalului de acceptare pentru probabilitățile de 1% cât şi 0,1% deci vom considera că diferențele sunt "extrem de semnificative", probabilitatea ca din întâmplare să obținem din aceeaşi populație două loturi atât de diferite fiind sub 0,1%.

Varianta II

Raționamentul expus mai sus este uşor de înțeles, fiind calculate intervalele de acceptare pentru 3 probabilități: 5%, 1% şi 0,1% şi urmărind încadrarea diferenței reale. În pachetele software de prelucrari statistice se procedează invers: se calculează direct probabilitatea de a obține asemenea diferențe din întamplare.

Se calculează mai întâi valoarea lui *t* care corespunde diferenței reale:

$$t_c = \frac{\bar{d}}{S_d} = \frac{\bar{X}_1 - \bar{X}_2}{S_d}$$

(II.5.5.)

adică:

$$t_c = \frac{0{,}56}{0{,}14} = 4.00$$

Din tabelul valorilor distribuției t, pentru 60 grade libertate (urmărim linia lui 60), vedem că $t_c > t_{0,999;\ 60}$; programele statistice ne dau valoarea pentru care:

$$t_c = t_{p,\nu} \qquad \text{(II.5.6.)}$$

şi afişează valoarea lui p, pe care o vom interpreta conform fig.II.14. În exemplul nostru obţinem *p = 0,00087*, deci având *p<0,001* vom spune că diferenţele sunt "extrem de semnificative".

Testul Mann-Whitney este echivalentul neparametric al testului **t** nepereche. Ca raţionament este similar cu testul Wilcoxon şi va fi exemplificat acolo. Pentru loturi mai mari, rezultatul obţinut este acelaşi ca în cazul aplicării testului **t.** Pachetele software de prelucrari statistice dau valoarea lui *p* (adică probabilitatea ca diferenţele observate în eşantionul de valori să fi apărut din întamplare, în condiţiile în care indivizii observaţi ar face pate dintr-o aceeaşi populaţie statistică). Interpretarea o facem tot conform fig. II.14.

C. Se compară **două valori medii,** din două serii obţinute pe **aceiaşi indivizi,** în două **condiţii diferite**

. Ipoteza zero: $H_0: \overline{X}_1 = \overline{X}_2$

. Condiţii: valori perechi - X_{i1} , X_{i2} reprezintă valorile obţinute pe individul **i** în condiţiile "1" respectiv "2".

Întotdeauna: $N_1 = N_2 = N$

Grade de libertate: $\nu = N - 1$

. Test aplicat: testul **t pereche**

Observaţii: testul **t** pereche este de fapt un test **t** pentru o serie aplicată diferenţelor; acest lucru este vizibil dacă sistematizăm datele conform tabelului II.6.

Tabel II.6. Prezentarea datelor pentru testul **t** pereche

Individ	**Valori experimentale** **Condiţia 1**	**Condiţia 2**	**Diferenţe** $D_i = X_{2i} - X_{1i}$
1	X_{11}	X_{12}	D_1
2	X_{21}	X_{22}	D_2
i	X_{i1}	X_{i2}	D_i
N	X_{N1}	X_{N2}	D_N
Medii	$\overline{X}_1$	$\overline{X}_2$	D

Valorile $\mathbf{d_i}$ pot fi pozitive sau negative; dacă între cele două serii nu sunt diferenţe vom avea $\overline{D} = 0$. Ipoteza de zero de mai poate deci scrie:

$$H_0 : \overline{D} = 0$$

Exemplu II.10. Dorim să studiem efectul unui medicament asupra frecvenţei cardiace. Pe un lot de *N = 9* indivizi obţinem valorile din tabelul II.7
Efectuând calculele, obţinem:

$$\overline{D} = +4,\ S = 4{,}5,\ S_{\bar{x}} = 1{,}5$$

Pentru $\nu = 8$ şi $\alpha = 5\%$, $t_t = 2{,}3$ deci regiunea de acceptare va fi:

$$(-\ t_t \cdot S_{\bar{d}}, +t_t \cdot S_{\bar{d}}) = (-2{,}3\ .\ 1{,}5\ ,\ +\ 2{,}3\ .\ 1{,}5) = (-3{,}45,\ +\ 3.45)$$

deci valoarea obținută $\overline{D}$ de găseşte înafara intervalului de acceptare şi vom spune că **diferențele sunt semnificative,** probabilitatea ca din întâmplare să obținem diferențele din tabelul II.53. fiind sub 5%.

Tabel II.7. Frecvența cardiacă înainte şi după tratament

Subiect	FC înainte	FC după	Diferența
1	***63***	***73***	***+10***
2	***67***	***67***	***0***
3	***79***	***76***	***-3***
4	***67***	***75***	***8***
5	***68***	***70***	***2***
6	***72***	***71***	***-1***
7	***73***	***80***	***7***
8	***69***	***76***	***7***
9	***70***	***75***	***5***

Introducând datele din tabelul II.53. într-un program de calculator obținem p = 0,042, adică $p < 0,05$, deci cu aceeaşi interpretare.

D. Se compară **două mediane sau două serii ordinale.**

. Ipoteza de zero se va referi la mediana ca indicator al tendinței centrale.

. Test aplicat: **testul Wilcoxon -** aplicat în două versiuni:

- pentru serii independente : testul "**suma rangurilor"** (*rank-sum test*)
- pentru serii perechi: testul "**semnul rangurilor"** (*signed-rank test*)

a.) Testul "**suma rangurilor"**

. Condiții: variabile ordinale (ranguri) sau variabile numerice aduse sub formă de ranguri (în acestă situație el este identic cu testul neparametric *Mann - Whitney*)

. Grade de libertate: valorile din tabel sunt dependente de ambele dimensiuni N_1 şi N_2; de aceea de acceptă să se ia convențional cea mai mică serie ca prima ($N_1 \leq N_2$)

. Fundamentare teoretică:

Fie două serii de valori ale unei variabile ordinale, obținută pe două loturi 1 şi 2 pe care le ordonăm astfel:

- seria 1, cu N_1 indivizi: $X_1 \leq X_2 \leq ..... \leq X_i \leq ..... \leq X_{N1}$
- seria 2, cu N_2 indivizi: $Y_1 \leq Y_2 \leq ..... \leq Y_j \leq ..... \leq Y_{N2}$

Seriile 1 şi 2 le luăm astfel încât $N_1 \leq N_2$.

Se alcătuieşte lotul compus din amestecarea celor două loturi, având *N = N1 + N2* indivizi şi se ordonează încât: $Z_1 \leq Z_2 \leq ..... \leq Z_k \leq ..... Z_N$, unde Z este o valoare X sau Y. Acestui şir i se asociază un şir de ranguri r_k cu valori între 1 şi N; dacă două sau mai multe valori succesive în şirul Z sunt egale (de exemplu $Z_2 = Z_3 = Z_4$), acestor ranguri li se asociază o valoare intermediară calculată ca medie între rangul maxim şi cel minim din acel grup de ranguri (în exemplul nostru $r_2 = r_3 = r_4 = 3.$) Să notăm deci rangurile cu: $r_1 \leq r_2 \leq .... \leq r_k \leq .... \leq r_N$ şi notăm suma rangurilor ce corespund valorilor din primul lot cu R_1 , respectiv din al doilea lot cu R_2 . Pentru aplicarea testului se calculează două statistici:

$$U_1 = N_1N_2 + N_1(N_1+1)/2 - R_1 \quad \text{(II.5.7.a)}$$

$$U_2 = N_1N_2 + N_2(N_2+1)/2 - R_2 \quad \text{(II.5.7.b)}$$

şi se ia $U = \min(U_1, U_2)$. (II.5.7.c)

Dacă N_1 şi N_2 sunt mari (> 10), statistica U are o distribuție aproximativ normală cu media:

$$\mu_U = N_1 \cdot N_2 / 2 \quad \text{(II.5.8)}$$

şi deviația standard:

$$\sigma_u = \sqrt{N_1N_2(N_1+N_2+1)/12} \quad \text{(II.5.9.a)}$$

Pentru eşantioane mai mici s-au realizat tabele speciale pentru testul *Wilcoxon rank-sum* ce conțin probabilitatea de a obține valori U în anumite inervale.

Cel mai des, din tabele se apreciază intervalul ce cuprinde regiunea de acceptare a ipotezei de zero cu o anumită probabilitate, adică regiunea de încadrare a valorii R_1 care are o repartiție cu media:

$$\overline{R}_1 = N_1(N_1+N_2+1)/2 \quad \text{(II.5.10)}$$

şi deviația standard

$$S_U = \sqrt{N_1N_2(N_1+N_2+1)/12} \quad \text{(II.5.9.b)}$$

Pentru un test bilateral cu α = 5%, se caută în tabele valorile pentru $R_1(0{,}025)$ şi $R_2(0{,}975)$.

Pachetele software de prelucrări statistice ne dau direct valoarea probabilității **p** interpretabilă conform fig. II.14.

Exemplu II.11. Se analizează aprecierea subiectivă a gradului de adaptare la efort al unui lot de sportivi comparativ cu un lot martor. Pentru aprecierea adaptării se foloseşte scara Borg a senzației subiective de efort, care asociază valori de la 0 la 20, aproximativ după tabelul II.54.

Tabelul II.8. Scara Borg a senzației subiective la efort (sumar)

Valoare	Aprecierea efortului
0	Extrem de uşor
5	Relativ uşor
10	Mediu
15	Destul de greu
20	Epuizant

Ambele loturi sunt supuse la un efort standard: 5 minute, 45 W efort triunghiular, pe bicicleta ergometrică. Un model de rezultate este prezentat în tabelul II.9.

Tabelul II.9. Senzaţia subiectivă la efort standard pe două loturi: lot 1 - sportivi (N_1=6), lot 2 - martor (N_2 =8); valorile sunt ordonate crescător pentru fiecare lot

Valori		**Ranguri**	
Sportivi	**Martor**	**Sportivi**	**Martor**
2	*6*	*1*	*5*
5	*6*	*2,5*	*5*
5	*9*	*2,5*	*8*
6	*11*	*5*	*10,5*
8	*11*	*7*	*10,5*
10	*12*	*9*	*12*
	14		*13*
	15		*14*
		R_1 =27	R_2 =78

$U_1 = 6 \cdot 8 + 6 \cdot 7 / 2 - 27 = 42$
$U_2 = 6 \cdot 8 + 8 \cdot 9 / 2 - 78 = 6$
$U = 6$
$\mu_u = 24$
$\sigma_u^2 = 60$

Sirul global	2	5	5	6	6	6	8	9	10	11	11	12	14	15
Nr. crt.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Ranguri	1	2.5	2.5	5	5	5	7	8	9	10.5	10.5	12	13	14
Lotul	1	1	1	(1)	(2)	(2)	1	2	1	2	2	2	2	2

Deoarece în cele două loturi sunt prea puţine valori pentru ca statistica U calculată să urmeze o distribuţie normală, vom determina nivelul de semnificaţie utilizând tabelele speciale pentru eşantioane mici.

Din tabelul pentru testul U se obţine, pentru un test bilateral, cu prag de semnificaţie
α =5% ($\alpha/2$ = 0,025 şi 1-$\alpha/2$ = 0,975), intervalul de acceptare pentru R_1: (29,61); observăm că valoarea noastră R_1 = 27 este în afara intervalului de acceptare a ipotezei zero, deci o respingem şi vom considera că sportivii apreciază acelaşi efort ca fiind mai uşor.

b.) Testul "**semnul rangurilor"**

. Condiţii: serii perechi de variabile ordinale (ranguri) sau variabile numerice aduse sub formă de ranguri (echivalentul neparametric pentru testul **t** pereche).

. Grade de libertate: $N_1 = N_2 = N$;

. Fundamentare teoretică:

Fie două serii de valori ale unei variabile ordinale, obţinute pe un lot, de volum N, în două condiţii, 1 şi 2. Pentru fiecare individ **i** obţinem diferenţa $D_i = X_{2i} - X_{1i}$, care poate fi pozitivă sau negativă. Se ignoră diferenţele nule.

Într-o primă fază neglijăm semnele şi ordonăm crescător valorile absolute ale diferenţelor; apoi le acordăm ranguri, ca în exemplul anterior. Reintroducem acum pentru ranguri semnele pe care le-am avut la diferenţele D_i şi calculăm separat două

totaluri: R (+) este suma rangurilor pozitive şi R(-) este suma rangurilor negative. Calculăm acum statistica testului:

$$R = R(+) \qquad \text{(II.5.11)}$$

$$T = \left[\left| R - \frac{N(N+1)}{4} \right| - \frac{1}{2} \right] \Big/ \sqrt{N(N+1)(2N+1)/24} \qquad \text{(II.5.11')}$$

N reprezintă numarul diferenţelor D_i care nu sunt zero.

Pentru eşantioane mari (N ≥ 16) statistica *T* are o repartiţie normală, cu media:

$$\mu_T = N(N+1)/4 \qquad \text{(II.5.12)}$$

şi deviaţia standard

$$\sigma_T = \sqrt{N(N+1)(2N+1)/24} \qquad \text{(II.5.13)}$$

Se caută din tabel valoarea $T_{0.025;10} = T_{tab}$ pentru a accepta H_0

Pentru situaţii în care statistica *T* nu urmează o distribuţie normală (numărul diferenţelor D_i care nu sunt zero N < 16), există tabele speciale care prezintă pentru testul *Wilcoxon signed-rank*. Se poate astfel evalua intervalul de acceptare a ipotezei zero, pentru testele bilaterale respectiv limitele critice pentru testele unilaterale.

Pachetele software de prelucrări statistice ne dau valoarea probabilităţii **p** de acceptare a ipotezei zero, interpretabilă conform fig. II.14.

Exemplu II.12. Aprecierea cunoştinţelor (aptitudinilor) prin note reprezintă variabile ordinale. Analizăm eficienţa unui curs după punctajul obţinut la un test de cunoştinţe aplicat atât înainte cât şi după un curs auxiliar pe acelaşi grup de studenţi. Rezultatul la un astfel de chestionar se exprima printr-un punctaj cu valori între 0 şi 40.

În tabelul II.10 sunt prezentate rezultatele obţinute pe un lot de 10 studenţi.

Tabelul II.10. Rezultatele la testul de biostatistică obţinute pe un lot de 10 studenţi, înainte şi după ce au urmat un curs auxiliar de teoria probabilităţilor

Student	Înainte	După	Diferenţa	Rang
1	35	38	+3	+5,5
2	26	30	+4	+7,5
3	36	36	0	
4	30	35	+5	+9
5	38	40	+2	+4
6	29	28	-1	-2
7	21	25	+4	+7,5
8	27	24	-3	-5,5
9	31	30	-1	-2
10	35	36	+1	+2

La stabilirea rangurilor se ignora diferenţele cu valoare 0.

Şir diferenţe absolute	0	1	1	1	2	3	3	4	4	5
Nr. crt.		1	2	3	4	5	6	7	8	9
Rang		2	2	2	4	5,5	5,5	7,5	7,5	9

R(+)=35.5

R(-)=9.5

În acest exemplu nu vom putea utiliza distribuţia normală (sunt prea putine valori), ci tabelul special pentru testul *Wilcoxon signed-rank*. Putem alege un test unilateral, adică ipoteza de zero să fie $H_0 : \Delta = 0$ (în cuvinte: cursul nu a determinat îmbunătăţirea semnificativă a rezultatelor la test), având ca alternativă, în cazul respingerii H_0 ipoteza $H_1 : \Delta > 0$ (adică rezultatele după cursul auxiliar sunt semnificativ mai bune). Din tabel, pentru $\alpha = 0{,}05$ şi $N = 9$ obţinem valoarea critică $R_{tab} = 40$. Pentru statistica T decizia se ia după regula:

Dacă $R_{calc} \rangle R_{tab}$, atunci respingem H_0. În cazul nostru concret nu vom putea respinge ipoteza de zero. Deci, în ciuda aparenţelor (la 6 din cei 10 nota a crescut şi a scăzut numai la 3, iar creşterile sunt mai mari decât scăderile), nu putem afirma că diferenţele sunt semnificative (probabilitatea să obţinem diferenţe de acest gen din întâmplare este destul de mare, peste 5%).

Acceptarea ipotezei de zero în acest caz nu înseamnă neapărat "cursul auxiliar nu a determinat creşteri semnificative ale rezultatelor" ci doar că "din analiza rezultatelor a 10 studenţi nu putem afirma existenţa unei creşteri semnificative a rezultatelor la testul de cunoştinţe". Deseori, în asemenea situaţii, când rezultatul unui test statistic este la limită, este bine să se extindă studiul pe un lot mai mare pentru a se putea atinge semnificaţia statistică.

Observaţie: Rezultatele obţinute ar fi asemănătoare şi dacă variabilele ar fi fost considerate numerice în loc de ordinale şi am fi aplicat testul **t** pereche pentru diferenţe având valoarea $\overline{D} = +1.4$ *şi* $S_{\overline{d}} = 0.83$ pentru (N = 10); concluziile ar fi şi în acest caz la fel ca cele obţinute cu testul Wilcoxon.

E. Se compară **n valori medii**

. Ipoteza de zero: $H_0 : \overline{X}_1 = \overline{X}_2 = \ldots = \overline{X}_n$ (II.5.14.)

. Test aplicat: ANOVA (ANalysis Of VAriance).

Testele de tip ANOVA formează o întreagă clasă ce face obiectul de studiu al unui întreg capitol al (bio) statisticii numit "analiza varianţei". Elementul central în analiza varianţei este împărţirea "varianţei" valorilor individuale (formula 5.7.b) în funcţie de originea posibilă (sursele) variaţiei.

Analiza varianţei se aplică pentru testarea egalităţii a **n** valori medii în două variante principale:

Analiza **unifactorială** ("one-way analysis") - pentru a compara **n serii independente,** obţinute pe loturi diferite.

Analiza **bifactorială** ("two-way analysis") - pentru a compara **n serii dependente,** obţinute pe acelaşi lot, în **n** condiţii.

a) Analiza unifactorială

. Tipuri de analiză unifactorială:

i^0 - parametrică: testul **t nepereche generalizat** pentru n serii

i^{00} - neparametrică: testul **Kruskal - Wallis**

. **Fundamentare teoretică:**

Fie **n** serii experimentale obţinute pe loturi independente, un lot **j** având N_j indivizi. Fiecare lot corespunde unei condiţii pe care o vom numi "tratament"; avem deci **n** tratamente; pentru tratamentul j analizăm lotul corespunzător.

X_{j1} , X_{j2} , ...X_{ji} , ..., X_{jNi} , având:

media $\overline{X}_j = (\sum_i^{N_j} X_{ji}) / N_j$ (II.5.15)

varianța $S_j^2 = SS_j / (N_j - 1)$ (II.5.16)

suma pătratelor abaterilor: $SS_j = \sum_i^{N_j} (\overline{X}_j - X_{ji})^2$; (SS = *sum of squares*).

Dacă amestecăm loturile obținem un grup mare având în total N indivizi, ale căror valori le notăm acum cu X_i :

$$N = \sum_j^n N_j \quad \text{(II.5.17)}$$

Pentru acest lot global avem o medie generală:

$$\overline{\overline{X}} = \left(\sum_{i=1}^{N} X_i \right) / N = \left(\sum_j^n N_j \overline{X}_j \right) / n \quad \text{(II.5.18)}$$

şi suma totală a pătratelor (TSS = *total sum of squares)*:

$$TSS = \sum_{i=1}^{N} (\overline{\overline{X}} - X_i)^2 \quad \text{(II.5.19)}$$

Esența în testele ANOVA este a diviza această **varianță totală TSS** (exprimată de fapt aici prin suma totală a abaterilor tuturor valorilor individuale X_i față de media globală $\overline{\overline{X}}$) în varianța datorată variațiilor valorilor individuale X_{ji} din fiecare grup *j* față de media grupului $\overline{X}_j$, numită **varianța reziduală** (RSS) şi **varianța datorită tratamentelor** (treatment variance).

$$TSS = RSS + TrSS \quad \text{(II.5.20)}$$

Suma pătratelor abaterilor pentru varianța reziduală este:

$$RSS = \sum_j^n SS_j = \sum_j^n (\sum_i^{N_j} (\overline{X}_j - X_{ji})^2)$$

(II.5.21) Pentru cele **n** serii (“tratamente”), cuprinzând un total de N indivizi , toate cele N valori sunt independente deci numărul gradelor de libertate (d.f. = *degrees of freedom*) pentru TSS este N. Numărul gradelor de libertate pentru tratamente este :

$$\nu = dfTr = n - 1 \quad \text{(II.5.22)}$$

iar pentru reziduale este:

$$\nu_2 = dfR = N - (n - 1). \quad \text{(II.5.23)}$$

De fapt RSS şi TrSS calculate cu (II.5.21) şi (II.5.20) sunt sume pătrate; pentru a reprezenta variante cu adevărat ele trebuiesc divizate cu numărul corespunzător de grade de libertate: (Ms = mean square).

$$MSTr = (TSS - RSS) / \nu_1 \qquad \text{(II.5.24)}$$

$$MSR = RSS / \nu_2 \qquad \text{(II.5.25)}$$

Raportul a două varianțe prezintă o distribuție F cu (ν_1;ν_2) grade de libertate:

$$F_{calc} = F = \frac{MSTr}{MSR}$$

Tabel II.11. Scăderile tensiunii sistolice după patru zile de tratament, pe 3 loturi. Valorile negative indică o creştere a tensiunii. În fiecare serie valorile au fost ordonate. Pentru seria globală avem:

Tratament j Individ i	1	2	3	
1 *2* *3*	*20* *15* *15*	*30* *25* *15*	*30* *25* *20*	
4 *5* *6*	*10* *5* *5*	*10* *10* *5*	*15* *10* -	
7 *8* *9* *10*	*0* *-5* *-5* *-10*	*5* *0* *-10* -	- - - -	
N_j	*10*	*9*	*5*	*N = 24*
$\overline{X}_j$	*5*	*10*	*20*	$\overline{\overline{X}}$ *=10*
SS_j	*900*	*1200*	*250*	*RSS =2350*
$\sum(\overline{\overline{X}} - X_i)^2$	*1150*	*1200*	*750*	*TSS = 3100*

Dacă seriile nu diferă semnificativ între ele, varianța care rămâne atribuită tratamentelor MSTr este mică, varianța totală fiind explicată aproape integral de rezidualele MSR (variațiile individuale din fiecare grup), deci F_{calc} va avea valori mici. Însă dacă seriile diferă semnificativ, MSTr va reprezenta o porțiune însemnată din varianța totală şi F va fi mare.

Stabilind un prag de semnificație α (5% sau 1%) decizia testului se ia astfel:

- dacă $F_{calc} > F^{\alpha}_{\gamma_1\gamma_2}$ (tabel), atunci $p < \alpha$, adică respingem H_0

- dacă $F_{calc} < F_{tab}$, atunci acceptăm H_0

Actualele pachete statistice prezintă rezultatele în forma standard a tabelelor ANOVA şi calculează direct valoarea lui **p** interpretabilă conform fig.II.14.

Exemplul II.13. Comparăm 3 tratamente antihipertensive obținând pe trei loturi rezultatele din tabelul II.11.

Din tabelul distribuției F avem:

$$F_{2,22}^{0,05} = 3{,}44 \quad \text{și} \quad F_{2,22}^{0,01} = 5{,}72$$

Cum $F_{calc} \rangle F_{2,22}^{0,05}$, rezultă $p < 0{,}05$ deci respingem ipoteza de zero H_0 vom considera că între cele 3 serii avem diferențe semnificative.

În tabelul II.12. sunt prezentate datele sintetice ale testului **ANOVA.**

Tabel II.12. Tabelul **ANOVA** cu datele brute din tabelul II.7

Sursa de variație	**Grade de libertate**	**Suma pătratelor SS**	**Media pătratelor Ms**	**Raportul F F**
Tratament Reziduale	2 22	750 2350	375 106,8	3,48
Total	24	3100		

Observații:

- Cel mai adesea, după aplicarea unui test ANOVA pentru mai mult de două serii, analiza poate continua prin compararea pe rând a câte două serii prin testul *t standard* (sau Mann-Whitney - Wilcoxon în caz neparametric) cu ajustarea corespunzatoare a lui α funcție de numarul de teste.
- Pentru numai două serii rezultatul obținut prin ANOVA este identic cu cel obținut prin testul *t nepereche*.

b) Analiza bifactorială

- Tipuri de analiză bifactorială:

i^0 - *parametrică*: testul **t pereche generalizat**
ii^{00} - *neparametrică*: testul **Friedman**

Fundamentare teoretică

Fie n serii de valori experimentale obținute pe același lot, cu volumul de N indivizi, în n condiții diferite. Fiecare serie de valori corespunde unei condiții pe care o vom numi și aici "tratament". Fiecare individ i este supus tuturor celor n tratamente. Spre deosebire de cazul anterior, când luam în considerare un singur factor ce ar putea influența varianța - tratamentul, de această dată vom lua în considerare și al doilea factor, de exemplu individul - fiecare individ are reacții particulare la fiecare tratament. În general în ANOVA bifactorială gruparea după primul factor se face în "tratamente"- j, iar după al doilea factor se face în "blocuri"- i.

Să facem următoarele notații:

X_{ji} - o valoare individuală pentru tratamentul *j* la blocul *i* (individul i)

$$\overline{X}_{j\bullet} = (\sum_{i}^{N} X_{ji}) / N \quad \text{- media unui tratament} \qquad \text{(II.5.27.a)}$$

$\bar{X}_{\bullet i} = (\sum_{j}^{n} X_{ji}) / n$ - media unui bloc (II.5.27.b)

$\bar{\bar{X}} = (\sum_{i}^{N}\sum_{j}^{n} X_{ji}) / (Nn) = (\sum_{i}^{N} \bar{X}_{\bullet i}) / N = (\sum_{j}^{n} X_{j\bullet}) / n$ - media global (II.5.27.c)

Gradele de libertate sunt:

. total: $\nu = N \cdot n$ (II.5.28.a)

. pentru tratamente: $\nu'_1 = n - 1$ (II.5.28.b)

. pentru blocuri: $\nu''_1 = N - 1$ (II.5.28.c)

. pentru reziduale: $\nu_2 = N * n - (N-1) - (n-1)$ (II.5.28.d)

Tabelul II.13. Prezentarea tabelului ANOVA pentru analiza bifactorială

Sursa variatiei	*Grade de libertate df*	*Suma patratelor SS*	*Media patratelor ms*	*Raportul F*
Tratamente	$\nu'_1 = n-1$	*TrSS*	$M'_1 = TrSS / \nu'_1$	$M'_1 / M_2(P')$
Blocuri (indivizi)	$\nu''_1 = N-1$	*BlSS*	$M''_1 = BlSS / \nu''_1$	$M''_1 / M_2(P'')$
Reziduale	$\nu_2 = N_n - (N+n-2)$	*RSS*	$M_2 = RSS / \nu_2$	
Total	$\nu = Nn$	*TSS*		

Sumele pătratelor vor fi calculate cu:

$$\text{TSS} = \sum_{i}^{N}\sum_{j}^{n} (\bar{\bar{X}} - X_{ij})^2 = \sum_{i}^{N}\sum_{j}^{n} X_{ij}^2 - Nn\bar{\bar{X}}^2 \quad \text{(II.5.29.a)}$$

$$\text{RSS} = \sum_{i}^{N}\sum_{j}^{n} ((\bar{X}_{j.} + \bar{X}_{.i} - \bar{\bar{X}}) - X_{ij})^2 \quad \text{(II.5.29.b)}$$

$$\text{TrSS} = \sum_{i}^{N}\sum_{j}^{n} (\bar{X}_{j.} - X_{ij})^2 = N\sum_{j} \bar{X}_{j.}^2 - Nn\bar{\bar{X}}^2 \quad \text{(II.5.29.c)}$$

$$\text{BlSS} = \sum_{i}^{N}\sum_{j}^{n} (\bar{X}_{.i} - X_{ij})^2 = n\sum_{i}^{N} \bar{X}_{.i}^2 - Nn\bar{\bar{X}}^2 \quad \text{(II.5.29.d)}$$

Între ele avem relația:

TSS = RSS + TrSS + BlSS (II.5.30)

Tabelul ANOVA pentru prezentarea rezultatelor va fi de forma II.13.

Pentru ca ordinea aplicării tratamentelor să nu fie aceeaşi la toţi indivizii (acesta ar putea influenţa efectele), se alocă tratamente în ordine întâmplătoare. Cel mai potrivit este să se folosească aşa numitul **pătrat latin** având pe linii sau coloane ordinea tratamentelor; de exemplu, pentru 4 condiţii sau tratamente (A,B,C,D) am putea avea pătratul:

$$M=\begin{vmatrix} A & B & C & D \\ C & A & D & B \\ D & C & B & A \\ B & D & A & C \end{vmatrix} \qquad \text{(II.5.31)}$$

Indivizii (sau blocurile) se atribuie întâmplător acestor succesiuni de tratamente.

F. Se compară **2 sau n dispersii** (deviații standard).

Uneori este necesară în practică verificarea egalității statistice a unor indicatori de dispersie - de exemplu, testele de semnificație aplicate pentru valori medii presupun o egalitate statistică a dispersiilor seriilor care trebuie testate înainte de aplicarea testului pentru medii.

Dintre testele pentru compararea indicatorilor de dispersie vom prezenta câteva mai des întâlnite:

- pentru a compara două deviații standard
- pentru a compara **n** deviații standard obținute pe serii diferite
- pentru a compara **n** deviații standard obținute pe același lot.

a. Se compară **două deviații standard**

. Ipoteza de zero: $H_0 : \sigma_1 = \sigma_2$

. Test aplicat: testul (exact) **F-Fischer-Snedecor**

. Fundamentare teoretică

Raportul a două dispersii ale unor populații cu distribuție normală prezintă o distribuție notată cu F, numită distribuția Fischer.

Fie două serii experimentale, de volume N_1 și N_2, având dispersiile S_1^2 și S_2^2; le notăm cu indicii 1 și 2 astfel încât $S_1^2 \geq S_2^2$ adică $S_{(1)} = \max(S_1, S_2)$. Se calculează raportul:

$$F = S_{(1)}^2 / S_{(2)}^2 \qquad \text{(II.5.32)}$$

Se alege tabelul cu valorile lui F după pragul de semnificație dorit (0,05 sau 0,01) și pentru cele două valori ale gradelor de libertate:

$$\nu_1 = N_1 - 1; \qquad \nu_2 = N_2 - 1 \qquad \text{(II.5.33)}$$

Ipoteza de zero se acceptă dacă indicele F respectă relația:

$$F_{calc} < F_{tab}^{\alpha/2}(\nu_1, \nu_2) \qquad \text{(II.5.34)}$$

Pachetele statistice dau de obicei valoarea lui **p** interpretabilă conform fig. II.14

Exemplu: Considerăm loturile 2 și 3 din tabelul II.11.

Pentru lotul 2: $S_2^2 = \frac{1200}{9-1} = 150 = S_{(1)}^2$

Similar: $S_3^2 = \frac{250}{5-1} = 62{,}5 = S_{(2)}^2$

Calculăm: $F = S_{(1)}^2 / S_{(2)}^2 = \frac{150}{62{,}5} = 2{,}4$. Din tabel, pentru $\alpha = 0{,}05$ avem: $F_{8,4}^{0,025} = 6{,}04$

deci $F_{calc} < F_{tabel}$ și vom admite H_0 deși diferențele dispersiilor păreau destul de mari.

b. Se compară **n deviații standard** obținute pe **serii diferite**

. Ipoteza de zero: $H_0: \sigma_1 = \sigma_2 = ... = \sigma_n$

. Test aplicat: testul lui **Bartlett**

. Fundamentare teoretică

Pentru n serii experimentale de volume N_j, medii $\overline{X}_j$ și deviații standard S_j, notăm:

- grade de libertate: $\nu_j = N_j - 1; \qquad \nu_i = \sum_{j=1}^{n} \nu_i = N - n$

(II.5.35)

- număr total de indivizi: $N = \sum_{j=1}^{n} N_j$ (II.5.36)

- dispersie globală: $S^2 = \frac{1}{\nu} \sum_{j}^{n} \nu_j S_j^2$ (II.5.37)

- coeficientul: $C = 1 + \frac{1}{3(n-1)} \left(\sum_{j}^{n} \frac{1}{\nu_j} - \frac{1}{\nu} \right)$ (II.5.38)

- statistica: $X_B^2 = \frac{1}{c} (\nu \cdot \ln S^2 - \sum_{j}^{n} \nu_j \ln S_j^2)$ (II.5.39)

este o variabilitate aleatoare cu distribuție χ^2 cu *n - 1* grade de libertate.

- regiunea de acceptare a ipotezei zero H_0 este dată de condiția:

X_B^2 calc $\leq \chi^2_{\alpha, n-1}(tab)$ (II.5.40)

și H_0 se respinge în caz contrar.

Pachetele software de prelucrari statistice ne dau direct valoarea lui **p.**

Exemplu: Să comparăm deviațiile standard ale celor 3 loturi prezentate în tabelul II.11

- seria 1: $N_1 = 10, \nu_1 = 9, S_1^2 = 900/9$
- seria 2: $N_2 = 9, \nu_2 = 8, S_2^2 = 1200/8$
- seria 3: $N_3 = 5, \nu_3 = 4, S_3^2 = 250/4$
- aplicăm formulele (II.5.36.) - (II.5.39.)
  - . $N = 24$, $n = 3$
  - . $S^2 = \frac{1}{24-3}(900 + 1200 + 250) = 112$
  - . $C = 1 + \frac{1}{3 \cdot 2} \left[\left(\frac{1}{9} + \frac{1}{8} + \frac{1}{4} \right) - \frac{1}{9+8+4} \right] = \frac{15}{14}$
  - . $X_B^2 = 9{,}59$
- din tabelul χ^2 pentru $\alpha = 0{,}05$ și *n-1 = 2* grade de libertate avem:

  $\chi^2_{0{,}05;2} = 5{,}99$

- observăm că $X_B^2(calc) \rangle \chi^2_{0{,}05;2}$, deci **respingem** H_0 și vom considera că dispersiile (derivațiile standard) diferă semnificativ.

c. Se compară **n deviații standard** obținute pe **aceiași indivizi**

. Ipoteza de zero: $H_0: \sigma_1 = \sigma_2 = ... = \sigma_n (sau \quad S_1 = S_2 = ... = S_n)$

. Test aplicat: testul lui **Cochran**

. Fundamentare teoretică:

Pentru *n* serii de date experimentale obținute pe același lot de volum *N*, având dispersiile $S_1, S_2, ..., S_n$, se alege:

$$S_{\max} = \max(S_1, S_2, ..., S_n) \qquad \text{(II.5.41)}$$

și se calculează:

$$S^2 = \sum_j^n S_j^2 \qquad \text{(II.5.42)}$$

Statistica

$$Q = S_{max}^2 / S^2 \qquad \text{(II.5.43)}$$

este comparată cu valoarea lui Q din tabelul lui Cohran; tabelele pentru Q sunt asemănătoare cu cele pentru F: sunt realizate pentru două valori ale lui α (0,05 și 0,01) și depind de 2 indici: *n* și $\nu = N - 1$
(numărul de grade de libertate).

Regiunea de acceptare a ipotezei de zero se alege dacă este satisfăcută condiția:

$$Q(calc) \langle Q_{n,\nu}^{\alpha}(tab) \qquad \text{(II.5.44)}$$

Pachetele software de prelucrari statistice dau direct valoarea lui **p** pentru interpretarea testului conform fig. II.14

G. se compară **proporții** (procente)

În cazul variabilelor nominale (calitative), indivizii din întregul eșantion sunt grupați în diferite clase, fiecare clasă având caracteristică o proporție (procent).

Dacă împărțirea se face în numai două clase, distribuția se numește **binominală,** dacă se face în mai multe clase se numește **multinomială.**

Testele pentru varibilele nominale sunt numeroase, acoperind toate categoriile de teste (semnificație, omogenitate etc.), astfel încât în cadrul cursului vom prezenta doar următoarele situații, mai des întâlnite:

a. se compară o proporție experimentală cu o valoare dată
b. se compară două proporții experimentale
c. se compară o distribuție experimentală cu una teoretică - test de concordanță
d. test de omogenitate pe tabel de contingență
e. test de independență pe tabel de contingență

a. Se compară o **proporție** experimentală **cu o valoare dată**

. Ipoteza de zero*:* $p = p_0$

. Condiție: se lucrează pe loturi mari, astfel încât să nu fie vreuna din clase cu mai puțin de 5 indivizi, deci N să fie altfel încât $N_{p0} \geq 5$ sau $N(1 - p_0) \geq 5$.

. Fundamentare teoretică

Statistica

$$Z = \frac{p - p_0}{\sqrt{\frac{p_0(1 - p_0)}{N}}} = \frac{D_p}{S_{p_0}} \qquad \text{(II.5.45)}$$

are o distribuție normală, deci intervalul de acceptare al ipotezei de zero va fi condiția:

$$Z_{(calc)} \in (-Z_{\alpha/2}, +Z_{\alpha/2}) \qquad \text{(II.5.46)}$$

adică pentru $\alpha = 5\%$ $\quad Z_{\alpha/2} = 1{,}96.$

Pachetele software de prelucrari statistice ne dau valoarea lui *p* (probabilitatea ca ipoteza de zero să fie adevărată) interpretabilă conform fig. II.14.

Exemplu II.14. Să se verifice dacă este adevărată afirmația că 4% dintre bărbați sunt daltoniști.

Cum $p_0 = 0{,}04$ şi trebuie ca $N_{p0} \geq 5$ rezultă $N \geq 125.$

Vom lua un lot de 150 bărbați pe care obținem $N_1 = 8$ daltoniști şi $N_2 = 142$ vedere colorată normală.

Avem deci $N = 150$, $\hat{p}_1 = 8 / 150 = 0{,}053$; $\hat{p}_2 = 142 / 150 = 0{,}946$

$$S_p = \sqrt{\frac{0.053 \cdot 0.946}{150}} = 0.018$$

$$Z = \frac{0{,}053 - 0{,}04}{0{,}01.6} = 0{,}833$$

Observăm că $Z_{calc} \in (-1.96; +1.96)$ deci **acceptăm** H_0**.**

b. Se compară **două proporții** experimentale

. Ipoteza de zero: $p_1 = p_2$

. Condiție: se lucrează pe loturi mari încât să nu fie vreuna din clase cu mai puțin de 5 indivizi

. Fundamentarea teoretică.

Statistica $Z = \dfrac{D_p}{S_{pd}} = \dfrac{p_1 - p_2}{S_{pd}}$ (II.5.47)

au o distribuție normală. Pentru eroarea standard a proporțiilor folosim formula (4.14.) sau

$$S_{pd} = \sqrt{p_0(1-p_0)\left(\frac{1}{N_1} + \frac{1}{N_2}\right)} \qquad \text{(II.5.48)}$$

unde:

$$p_0 = \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2} \qquad \text{(II.5.48.b)}$$

Regiunea de acceptare este dată tot de (II.5.46.).

Exemplu II.15. Verificăm dacă proporția de decese prin cancer de plămâni este aceeaşi la bărbați şi femei. Într-un studiu efectuat pe două loturi:

- lot 1: $N_1 = 180$ certificate deces bărbați înregistrate în 3 luni, din care $p_1 = 14 / 180 = 0{,}0777$ cu diagnosticul de mai sus
- lot 2: $N_2 = 165$ certificate deces femei, acelaşi interval, cu $p_2 = 5 / 165 = 0{,}030$

$$S_{pd} = \sqrt{\frac{0{,}077 \cdot 0{,}9222}{180} + \frac{0{,}03 \cdot 0{,}97}{165}} = 0{,}024$$

$$Z = \frac{(0{,}0777 - 0{,}030):100}{2{,}4} = 1{,}98$$

Observăm că $Z_{(calc)} \rangle Z_{\alpha/2} = 1{,}96$ deci **respingem** H_0 şi vom considera că decesul prin cancer de plămâni apare semnificativ mai frecvent la bărbaţi decât la femei.

c. Se compară o **distribuţie experimentală** cu o **distribuţie teoretică** de proporţii

În cazul variabilelor nominale, dacă avem **n** clase (distribuţie multinominală), rezultatele experimentale obţinute prin analiza unui lot de N indivizi se exprimă cu ajutorul unui tabel de forma tabelului II.14.

. Ipoteza de zero: $O_j = E_j \quad \forall j = 1, \ldots, n$ adică valorile observate experimental O_j diferă semnificativ de cele aşteptate E_j ("observed values"), pentru toate clasele j.

Tabelul II.14. Prezentarea datelor pentru aplicarea testului de concordonţă la o distribuţie multinominală.

Caracteristica	Clasa 1	Clasa 2	.....	Clasa n	Total
Valori experimentale	O_1	O_2	.....	O_n	N
Valori teoretice	E_1	E_2		E_n	N

. Test aplicat: **testul** χ^2 al lui **Pearson**

. Fundamentare teoretică:

Statistica

$$\chi^2 = \sum_j^n \frac{(O_i - E_i)^2}{E_i} \qquad \text{(II.5.49)}$$

are o repartiţie χ^2. Pentru ca ipoteza de zero să poată fi respinsă:

$$\chi^2(calc) \le \chi^2_{\alpha,\nu}(tab) \qquad \text{(II.5.50)}$$

. Pachetele statistice dau valoarea lui p interpretabilă conform fig. II.14.

. ***Observaţie:*** Valorile teoretice trebuiesc calculate în funcţie de specificul studiului; ele pot fi şi valori fracţionare. Ele pot fi evaluate şi pentru un total diferit şi apoi convertite pentru acelaşi total.

Exemplul II.16. Dorim să studiem răspândirea grupelor sanguine şi facem ipoteza că sunt uniform răspândite. Rezultatele experimentale obţinute pe un lot de 80 de indivizi sunt prezentate în tabelul II.15.a.

Tabel II.15.a. Repartiţia grupelor sanguine într-un lot cu *N = 80*

(Caracteristica) Grupa sanguină	O(I)	A(II)	B(III)	AB (IV)	Total
Valori experimentale	22	33	14	11	80
Valori teoretice	20	20	20	20	80

Ipoteza de zero: grupele sanguine sunt uniform răspândite în populaţia analizată.
Conform formulei (6.49.) obţinem:

$$\chi^2 = \frac{(22-20)^2}{20} + \frac{(33-20)^2}{20} + \frac{(14-20)^2}{20} + \frac{(11-20)^2}{20} = 14.5$$

Din tabel, pentru $\alpha = 5\%$ găsim $x^2_{0,05;3} = 7{,}815$ deci

$\chi^2(calc) \langle \chi^2(tab)$ şi vom respinge H_0 afirmând că din studiul efectuat rezultă că grupele sanguine nu au o răspândire uniformă în populaţia analizată.

Observaţie: noi am efectuat calcule anterioare pentru ipoteza că grupele sanguine ar fi uniform răspândite. Putem însă să verificăm şi alte ipoteze. De exemplu, un studiu efectuat în America Latină afirmă că acolo grupele sanguine ar avea răspândirea: 30% grupa 0, 15% grupa A, 40% grupa B şi 15% grupa AB. În acest caz tabelul II.5.11. ar deveni:

Tabel II.5.11. Repartiţia grupelor sanguine într-un lot cu *N = 80*

(Caracteristica) Grupa sanguină	O(I)	A(II)	B(III)	AB (IV)	Total
Valori experimentale	22	33	14	11	80
Valori teoretice	24	12	32	12	80

iar $\chi^2 = 15.9$ deci şi in această situatie ipoteza H_0 este respinsă.

d. Test de omogenitate pentru tabele de contingenţă

Să introducem mai întâi noţiunea de tabel de contingenţă. În cazul variabilelor nomiale (sau variabile numerice dar cu valori împărţite pe intervale), dacă urmărim impărţirea după două caracteristici (două criterii de clasificare) obţinem un tabel de contingenţă.

Definiţie: Tabelul de contingenţă reprezintă o formă de prezentare a datelor variabilelor nominale (sau pe clase) după două caracteristici: una plasată pe linii şi alta plasată pe coloane.

Un exemplu de tabel de contingenţă este prezentat în tabelul II.5.12.

Tabel II.16. Model de tabel de contingenţă

		elev/student	*muncitor*	*ţăran*	*intelectual*	*alte*	*Total*
Primul criteriu de clasificare: Mediul	*Urban*						
	Rural						

Tabelul expus se numeşte tabel 2 x 5 arătând numărul de clase după primul, respectiv al doilea criteriu de clasificare.

Un **test de omogenitate** aplicat unui **tabel de contingenţă** are menirea de a verifica dacă proporţiile diferitelor clase pe un rând (coloană) sunt aproximativ aceleaşi şi pe celelalte rânduri (coloane).

Valorile experimentale într-un tabel de contingenţă le notăm ca în tabelul II.17.

Tabel II-17. Notații în tabelul de contingență cu L linii şi C coloane pentru un eşantion de N indivizi

Crit. 2 **Crit. 1**	**1 ..... j ..... C**	**Total**
1 .... i L	σ_1 ..... σ_{1j} ..... σ_{1C} .... σ_{i1} ..... σ_{ij} ..... σ_{ic} .... σ_{L1} .... σ_{Lj} ..... σ_{L*}	L_{1*} ... L_{i*} .... L_{i*}
Total	C_{*1} .... C_{*j} ..... C_{*c}	**N**

Pentru notațiile din tabelul II.17 sunt adevărate relațiile de mai jos:

$$C_{*j} = \sum_{i}^{L} \sigma_{ij} \qquad \text{(II.5.51.a)}$$

$$L_{i*} = \sum_{j}^{C} \sigma_{ij} \qquad \text{(II.5.51.b)}$$

$$N = \sum_{j}^{C} C_{*j} = \sum_{i}^{L} L_{i*} = \sum_{i}^{L}\sum_{j}^{C} \sigma_{ij} \qquad \text{(II.5.51.c)}$$

$$\nu = (L-1)\,(C-1) \qquad \text{(II.5.52)}$$

Valorile aşteptate E_{ij} se calculează pentru fiecare element al tabelului după relația:

$$E_{ij} = \frac{L_{i*} \cdot C_{*j}}{N} \qquad \text{(II.5.53)}$$

astfel încât totalurile pe linii şi coloane vor rămâne nemodificate. De obicei se construieşte încă un tabel de forma tabelului II.7 cu deosebirea că, în loc de valorile observate, în căsuțe se trec valorile aşteptate. Marginile vor rămâne nemodificate.

Având ambele tabele putem trece la aplicarea testului.

. Ipoteza de zero : $H_0 : O_{ij} = E_{ij}$, pentru $\forall_{i,j}$

. Test aplicat : testul χ^2 al lui **Pearson.**

. Fundamentare teoretică:

Statistica

$$X^2 = \sum_i \sum_j \frac{(\sigma_{ij} - E_{ij})^2}{E_{ij}} \qquad \text{(II.5.54.)}$$

are o distribuție χ^2 cu ν grade de libertate.

Pentru un prag de semnificație ales, α, se caută în tabel valoarea $X^2_{\alpha,\gamma}$ şi se stabileşte regiunea de acceptare când este îndeplinită condiția (6.50.).

Programele de calculator dau direct valoarea lui p.

Exemplul II.17.a. Comparăm două tratamente A şi B pe două loturi având $N_1 = 100$ respectiv $N_2 = 50$ pacienți. Rezultatele obținute sunt trecute în tabelul II.17.a.

Tabelul II.17.a. Rezultatele obținute prin două tratamente

	Ameliorat	***Neameliorat***	***Total***
Tratament. A	40	60	100
Tratament B	30	20	50
Total	70	80	150

. Ipoteza de zero: H_0: cele două tratamente dau rezultate identice:
. Calculăm valorile aşteptate conform (6.53.) şi obținem tabelul II.17.b.

Tabelul II.17.b. Rezultate aşteptate în condiția respectării ipotezei de zero

	Ameliorat	***Neameliorat***	***Total***
Tratament. A	46,6	53,3	100
Tratament B	23,3	26,6	50
Total	70	80	150

Cu formula (5.54) obținem: $\chi^2 = 5{,}61$, în timp ce din tabel, pentru $\alpha = 5\%$ *şi* $\nu =$ *1* grad de libertate (pentru un tabel 2 x 2, cu 2 linii şi 2 coloane), avem $x^2_{0{,}05;1} = 3{,}84$.

Observăm că nu se respectă condiția $X^2(calc) \le X^2(tab)$, deci **nu acceptăm** ipoteza de zero şi vom spune că tratamentele dau rezultate diferite.

e. Test de independență pentru tabele de contingență

Un test de independeță are scopul de a stabili dacă există vreo relație de dependență între categoriile obținute prin două clasificări diferite; de exemplu între culoarea părului şi sex, între înălțime şi greutate, între vârstă şi adaptarea la efort etc.
. Ipoteza zero: cele două criterii de clasificare sunt independente (din punct de vedere probabilistic).
. Test aplicat: **testul χ^2 al lui Pearson**
. Fundamentare teoretică

Abordarea este asemănătoare cu cea prezentată la testul χ^2 ca test de omogenitate, având şi aceleaşi criterii de interpretare pentru regiunea de acceptare / respingere a H_0.

Exemplu II.18. Pentru a stabili dacă între înălțime şi greutate există vreo dependență , în cea mai simplă variantă putem alege o valoare care împarte, în două categorii aproximativ egale o populație din care extragem un eşantion. Rezultatele culese sunt prezentate în tabelul II.18.a.

Tabelul II.18.a. Clasificarea indivizilor unui lot după înălțime și greutate

Înălțime Greutate	sub 175 cm	peste 175 cm	Total
sub 70 kg	40	16	56
peste 70 kg	8	36	44
Total	48	52	100

. Ipoteza de zero: cele două clasificări sunt independente.
. Conform ipotezei de zero putem calcula valorile aşteptate cu formula (6.53) şi obținem tabelul II.18.b.

Tabelul II.18.b Valorile aşteptate la clasificarea indivizilor după înălțime şi greutate dacă cele două clasificări ar fi independente

Înălțime Greutate	sub 175 cm	peste 175 cm	Total
sub 70 kg	26,88	29,12	56
peste 70 kg	21,12	22,88	44
Total	48	52	100

Cu formula (5.54.) obținem $\chi^2 = 27{,}5$, în timp ce din tabel, pentru $\alpha = 5\%$ şi $\nu = 1$ grad de libertate avem $\chi^2_{0,05;1} = 3{,}84$; mai mult, chiar pentru $\alpha = 1\%$ şi $\gamma = 1$ avem $\chi^2_{0,005;1} = 7{,}88$ deci putem **respinge** H_0 şi să spunem că diferențele sunt foarte semnificative. Respingerea lui H_0 în cazul nostru înseamnă respingerea ipotezei că cele două criterii de clasificare sunt independente. O analiză mai detaliată a dependenței între variabile se face prin metode adecvate ce vor fi prezentate în capitolul următor.

Observație: În cazul tabelelor de contingență, dacă vreuna din căsuțe (mai ales ale valorilor aşteptate) conține mai puțin de 5% din observații, se preferă aşa numita "corecție Yates":

$$\chi^2 = \sum_i \frac{(/O_i - E_i / - 0{,}5)^2}{E_i} \qquad \text{(II.5.55.)}$$

astfel încât valorile foarte scăzute (mai rare) să nu influențeze prea puternic valoarea testului.

6. CORELAȚIA SI REGRESIA

După cum am sesizat încă din cursul precedent, mărimile pe care le analizăm în diferite studii pot fi, fie independente între ele, fie legate prin diferite relații. Evidențierea unor relații între mărimi poate sugera fie o fenomenologie cauzală, fie o corelație mai complexă ce necesită studii aprofundate. Oricum, respingerea unei ipoteze de zero într-un test de independență dă în general de gândit cercetătorilor, care pot sesiza o serie de aspecte interesante din simpla analiză statistică a datelor. Deseori analiza statistică a unor date sugerează o serie de alte studii pentru precizarea fenomenelor care generează anumite dependențe.

Datorită faptului că analizele de acest tip urmăresc comportarea a două variabile ele se numesc **analize bivariate.**

6.1. RELAȚII ÎNTRE DOUĂ VARIABILE CANTITATIVE

A. Relația de dependență

a. Variabile independente

Variabilele cantitative, fiind foarte des întâlnite în studiile biomedicale, permit cea mai fină analiză a relației de dependență/independență. Să ilustrăm acestea prin câteva exemple.

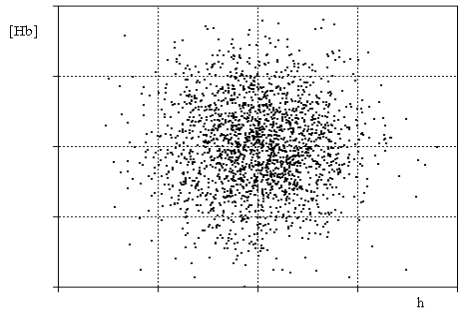


Figura II.16. Relația între înălțimea unui individ h, şi concentrația de hemoglobină din sânge [Hb]. Repartiția aproape simetrică şi uniformă a punctelor sugerează absența vreunei corelații

Exemplul II.19. Într-un studiu pe un lot de 50 de indivizi am urmărit mai multe variabile, cantitative şi calitative. Dacă alegem două variabile (cantitative), de exemplu înălțimea h, respectiv concentrația hemoglobinei în sânge, [Hb], într-o reprezentare grafică în care luăm pe axa Ox înălțimea h şi pe pe axa Oy concentrația hemoglobinei [Hb], fiecare individ va fi reprezentat print-un punct. Un astfel de grafic se numeşte "grafic de împrăştiere" ("scatter plot"). Datele obținute sunt reprezentate în figura II.16.

Repartiția simetrică şi fără vreo tendință a punctelor în graficul obținut sugerează absența vreunei legături între cele două mărimi; vom spune despre ele că sunt **independente.**

b. Variabile dependente

Dacă reprezentăm, relația între presiunea parțială a oxigenului din aerul respirat şi concentrația oxigenului dizolvat în sânge am obține un grafic de forma celui din fig. II.17.

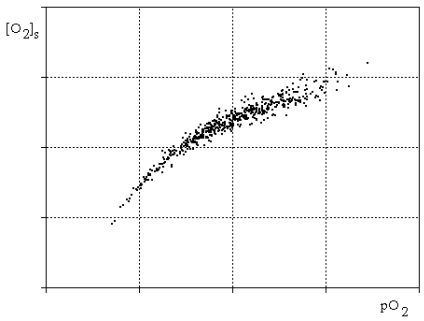


Figura II.17. Dependeța concentrației sanguine a oxigenului dizolvat de presiunea parțială a oxigenului din aerul respirat

Legătura dintre cele două mărimi este atât de vizibilă încât ne sugerează nu numai acceptarea unei relaţii cauzale ci chiar găsirea unei formule pentru relaţia dintre cele două mărimi; stabilirea unei astfel de formule ("formalizarea" matematică a fenomenului) reprezintă obiectul de studiu al unui capitol important al informaticii medicale numit "modelare şi simulare". În partea de biostatistică ne interesează doar faptul că cele două mărimi nu par independente - la testul χ^2 de independenţă, împărţind pO_2 şi respectiv $[O_2]$ în câteva clase (chiar şi cu numai 2 clase), vom respinge ipoteza de zero referitoare la independenţă, iar acceptarea unei dependenţe ne impinge spre căutarea unei formule care să exprime respectiva dependentă.

c. Variabile corelate

În exemplul prezentat anterior, relaţia cauzală părea rezonabilă: în condiţiile unei concentraţii crescute a oxigenului atmosferic pare uşor acceptabilă (cauzal) o concentraţie mai mare a oxigenului dizolvat în sânge. Există însă situaţii în care datele experimentale sugerează o relaţie de dependenţă, dar mecanismele fiziologice, la nivelul cunoştinţelor actuale, nu justifică pe deplin o relaţie cauzală directă, însă cel mai adesea admite o cauză comună pentru variaţiile observate ale celor două mărimi; astfel de variabile se numesc **variabile corelate**.

Un exemplu tipic îl constituie corelaţia între înălţimea şi greutatea indivizilor (figura II.18.), în care nu putem preciza că una dintre variabile este cauza şi cealaltă este efectul.

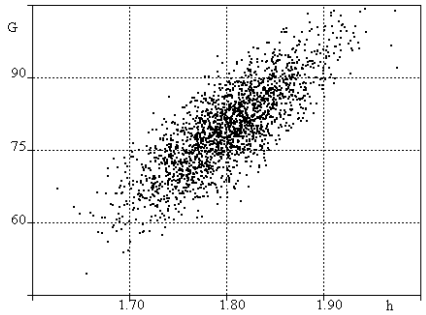


Figura II.18 Corelaţia înălţime-greutate pe un lot de 180 bărbaţi adulţi

Analiza corelaţiei înălţime-greutate, pe care o putem efectua când cunoaştem poziţia fiecărui punct în graficul II.18. este mult mai fină decât cea din testul de independenţă din cursul precedent. Repartizarea punctelor în graficul din figură, sugerează o exprimare de forma "cu cât individul este mai înalt, cu atât greutatea sa ne aşteptăm să fie mai mare."

B. Corelaţia liniară

În cazul în care considerăm că punctele dintr-o diagramă de împrăştiere se situează pe o dreaptă, corelaţia se numeşte **corelaţie liniară.**

a. Coeficient de corelaţie

"Intensitatea" corelaţiei este apreciată print-un parametru numit **coeficient de corelaţie Pearson.**

i ^{0} **- Formula** coeficientului de corelaţie este:

$$r = r_{xy} = \frac{s_{xy}}{S_x S_y} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2} \cdot \sqrt{\sum (Y_i - \bar{Y})^2}} \qquad \text{(II.6.1)}$$

unde

S_x^2 si S_y^2 reprezintă varianţa lui x, respectiv y:

$$S_x^2 = \frac{\sum (X_i - \bar{X})^2}{N} \quad , \; S_y^2 = \frac{\sum (Y_i - \bar{Y})^2}{N} \qquad \text{(II.6.2)}$$

iar S_{xy} se numeşte convarianţa între x şi y şi este dat de:

$$S_{xy} = \frac{\sum (X_i - \bar{X})\,(Y_i - \bar{Y})}{N} \qquad \text{(II.6.3)}$$

ii ^{0} **- Proprietăţi.**

- coeficientul de corelaţie r are valori cuprinse între -1 şi +1

$$r \in [-1, +1] \qquad \text{(II.6.4)}$$

- valorile pozitive ale lui r indică o corelaţie directă între x şi y (creşterea lui x este însoţită de creşterea lui y, figura II.19.a), în timp ce valori negative indică o corelaţie inversă (când x creşte, y scade, figura II.19.b.).

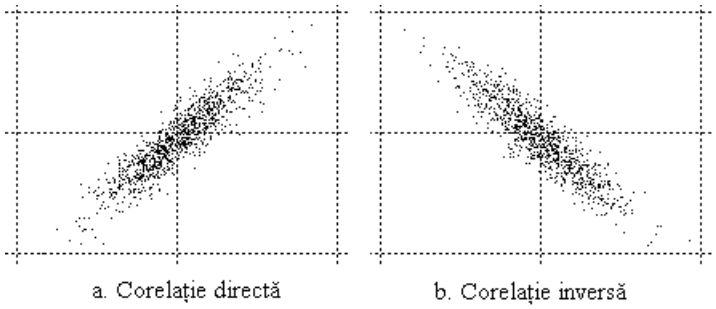


Figura II.19. Ilustrarea corelaţiei liniare directe şi inverse

- Valorile absolute mari ale lui r (apropiate de +1, respectiv -1) indică o corelaţie puternică, în timp ce valorile mici (în jurul lui 0) indică o corelaţie slabă (sau absenţa corelaţiei) - figura II.20.

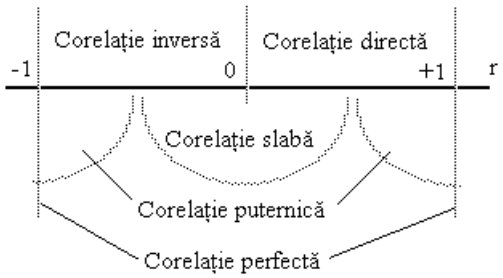


Figura II.20. Aprecierea "intensităţii" corelaţiei liniare după valoarea lui r

Observaţie: Coeficientul de corelaţie Pearson arată numai în ce măsură datele experimentale se potrivesc unei reprezentări descrise de o dreaptă; deci o valoare scăzută a lui r nu înseamnă neapărat corelaţie slabă ci *corelaţie liniară slabă*, însă ar putea fi puternică dar de alt tip.

b. Semnificaţia coeficientului de corelaţie

Valorile lui r depind atât de gradul de împrăştiere al valorilor experimentale cât şi de N - numărul de puncte. Uneori, când N este mic putem obţine, din întâmplare, valori ridicate pentru r, conducându-ne la concluzii hazardate cu privire la intensitatea corelaţiei. De aceea, se poate testa semnificaţia coeficientului de corelaţie liniară r.

. Ipoteza de zero: $H_0: \rho = 0$ ($\rho =$ coeficientul de corelaţie liniară pentru întreaga populaţie, r = coeficientul de corelaţie obţinut pe un eşantion).

. Test aplicat: testul t (**Student)**

. Fundamentare teoretică:

Se poate demonstra că raportul:

$$t = t_{calc} = r \cdot \sqrt{\frac{N-2}{1-r^2}} \qquad \text{(II.6.4)}$$

are o repartiţie Student cu $\nu = N - 2$ grade libertate.

Pentru un prag de semnificaţie α găsim în tabel valoarea $t_{\alpha/2,\nu}$. În caz că $t_{calc} \leq t_{tab}$ vom accepta H_0; în caz contrar o respingem şi vom spune că avem o probabilitate ridicată de a avea într-adevăr o corelaţie liniară.

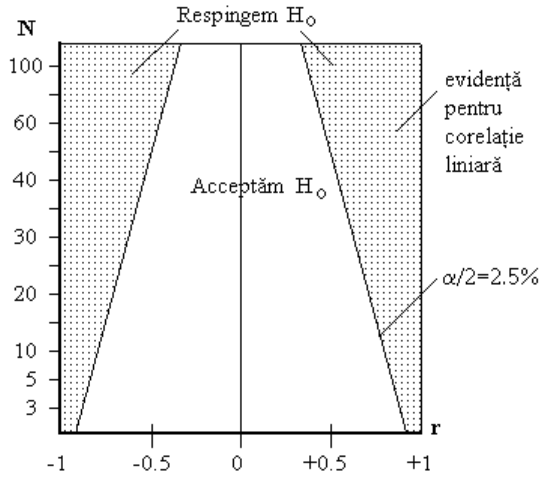


Figura II.21. Regiunile de acceptare/respingere a H_0 pentru coeficientul de corelaţie.

Eroarea standard a coeficientului de corelaţie se calculează cu relaţia:

$$S_r = \frac{1-r^2}{\sqrt{N}} \qquad \text{(II.6.5)}$$

deci pentru pragul de semnificaţie α , putem localiza intervalul în care se găseşte coeficientul de corelaţie al populaţiei ρ prin relaţia:

$$\hat{r} = \rho \in (r - t_{\alpha/2,\gamma} \cdot S_r \; ; \; r + t_{\alpha/2,\nu} \cdot S_r) \qquad \text{(II.6.6)}$$

Pe baza relației (II.6.6) se poate construi un tabel sau se poate ridica un grafic cu regiunea de acceptare / respingere a ipotezei de zero (figura II.21).

c. Dreapta de regresie

i ^{0} **- Definiție**

În cazul unei corelații liniare, dreapta care trece "cel mai bine" printre punctele experimentale se numeşte **dreaptă de regresie.**

ii ^{0} **- Ecuația dreptei de regresie**

Dacă notăm cu **x** variabila independentă şi cu y variabila dependentă, atunci ecuația unei drepte $y = f(x)$ are forma:

$$y = a + bx \qquad \text{(II.6.7)}$$

în care a se numeşte **ordonată la origine** (limba engleza "*intercept*") iar b se numeşte **panta** dreptei (limba engleza "*slope*") - figura II.22.

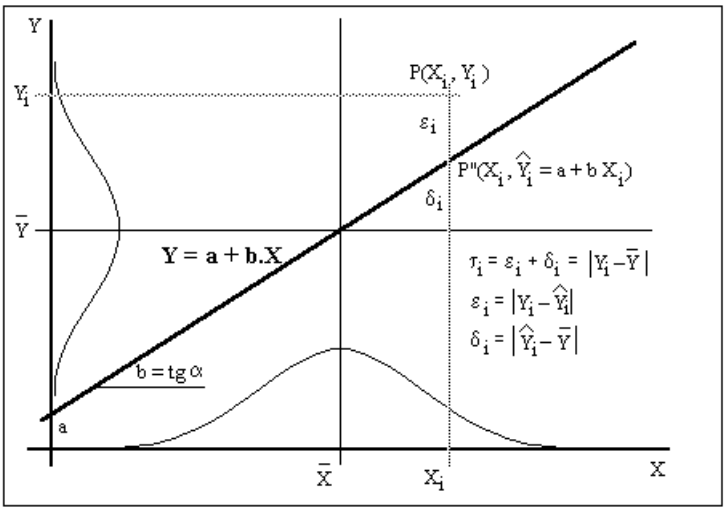


Figura II.22. Semnificația parametrilor pentru dreapta de regresie liniară.

iii ^{0} **- Metoda celor mai mici pătrate.**

Pentru determinarea coeficienților a şi b din ecuația dreptei vom considera că "cea mai bună" dreaptă care trece printre punctele experimentale este cea pentru care "suma pătratelor abaterilor, ε_i este minimă" , adică:

$$SSE = \sum \varepsilon_i^2 = min. \qquad \text{(II.6.8)}$$

iv ^{0} **- Formule pentru coeficienții dreptei de regresie.**

. Fundamentare teoretică.

Observăm că pentru un punct experimental $P(X_i, Y_i)$, găsim dreapta de regresie punctul $P'(X_i, \hat{Y}_i)$ la distanta $\varepsilon_i = Y_i - \hat{Y}_i$ unde $\hat{Y}_i$ reprezintă valoarea pe care ar avea-o variabila Y pentru valoarea lui dacă punctul s-ar găsi pe dreaptă:

$$\hat{Y}_i = a + bx_i$$

(II.6.9.)

Suma SSE depinde de coeficienții a şi b:

$$SSE = \sum (y_i - a - bx_i)^2 = \min \qquad \text{(II.6.8.b)}$$

Valoarea minimă se obține când derivatele în raport cu a şi b se anulează:

$$\frac{\partial SSE}{\partial a} = 0\ ,\ \frac{\partial SSE}{\partial b} = 0 \qquad \text{(II.6.8.c)}$$

Se obține un sistem de două ecuații cu două necunoscute, a şi b, care prin rezolvare ne dă rezultatele:

$$b = \frac{S_{xy}}{S_x^2} = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2} = Y_{xy}\frac{S_y}{S_x} \qquad \text{(II.6.9.a)}$$

$$a = \overline{Y} - b \cdot \overline{X} \qquad \text{(II.6.9.b)}$$

v^0 - Intervale de încredere pentru a şi b

. Fundamentarea teoretică.

Celor N perechi de valori reprezentând cele N puncte li se asociază $\nu = N - 2$ grade de libertate, ele fiind legate şi prin relația dreptei de regresie. Dacă notăm dispersia abaterilor cu S^2:

$$S^2 = \frac{SSE}{N-2} = \frac{\sum \varepsilon_i^2}{N-2} = \frac{\sum (y_i - \hat{y})^2}{N-2} \qquad \text{(II.6.10)}$$

atunci eroarea standard pentru pantă este:

$$S_b = \frac{\sqrt{S^2}}{\sqrt{S_x^2}} \qquad \text{(II.6.11.a)}$$

iar pentru ordonata la origine

$$S_a = \sqrt{S^2} \cdot \sqrt{\frac{1}{N} + \frac{X^2}{\sum (X_i - \overline{X}^2}} \qquad \text{(II.6.11.b)}$$

Pentru un prag de semnificație α , intervalele de localizare ale parametrilor estimați, $\hat{a}$ si $\hat{b}$ vor fi date de:

$$\hat{b} \in (b - t_{\alpha/2,\nu} \cdot S_b, b + t_{\alpha/2,\nu} \cdot S_b) \qquad \text{(II.6.12.a)}$$

$$\hat{a} \in (a - t_{\alpha/2,\nu} \cdot S_a, a + t_{\alpha/2,\nu} \cdot S_a) \qquad \text{(II.6.12.b)}$$

i^0 - Teste de semnificație pentru a şi b.

Având calculate intervalele de încredere a estimatorilor, putem aplica teste de semnificație pentru cei doi coeficienți ai dreptei de regresie.

. Pentru pantă:

- ipoteza de zero: $H_0 : b = 0$

- test aplicat: **testul t** pentru un prag de semnificație α ales şi pentru $\nu = N - 2$ grade de libertate, din tabel avem $t_{\alpha/2,\gamma}$. Calculăm:

$$t_{calc}^b = b / s_b \qquad \text{(II.6.13.a)}$$

dacă $t_{calc} \le t_{tab}$ se acceptă H_0, în caz contrar se respinge.

. Pentru ordonata de origine

- ipoteza de zero : $H_0 : a = 0$

- test aplicat: **testul t** pentru un prag de semnificație α ales şi pentru $\nu = N - 2$ grade de libertate, din tabel avem $t_{\alpha/2,\nu}$. Calculăm:

$$t^a_{calc} = a / S_a \qquad \text{(II.6.13.b)}$$

şi dacă $t_{calc} \leq t_{tab}$ se acceptă H_0, în caz contrar se respinge.

Observație: aplicarea testului de semnificație pentru pantă este foarte importantă deoarece o valoare nesemnificativ diferită de zero arată o independență între variabile, **chiar dacă** este mare şi satisface testul de semnificație.

Exemplul II.20. Corelația între *pH-ul* sanguin şi frecvența cardiacă poate fi reprezentată grafic ca în figura II.23.

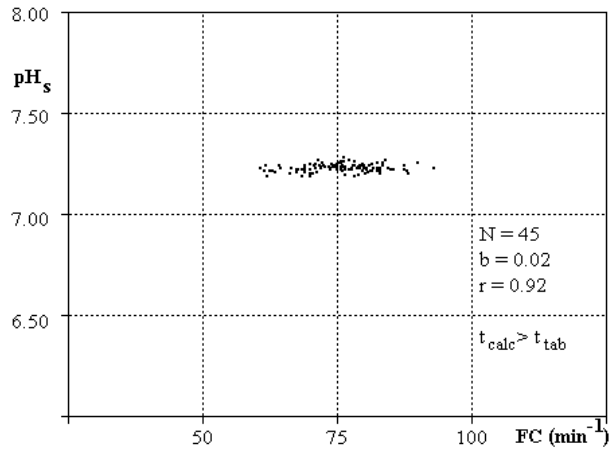


Figura II.23. Corelația între pH-ul sanguin şi frecvența cardiacă: deoarece pHs variază foarte puțin, punctele se înscriu bine pe o dreaptă şi obținem o valoare mare pentru r, care satisface şi testul de semnificație t. Însă panta nu diferă semnificativ de zero, deci putem considera mărimile ca independente

vii^0 **- Originea denumirii** dreptei "de regresie"

Numele de dreaptă de regresie a fost introdus de W. Galton, care a studiat relația între înălțimea copiilor şi înălțimea părinților. Deşi pe ansamblu copiii au avut o înălțime medie mai ridicată decât a părinților, această creştere nu era uniform repartizată, fiind mai accentuată pentru copiii având părinți mai scunzi, în timp ce înălțimea copiilor provenind din părinți înalți era deseori mai mică decât a părinților (figura II.24.).

Interpretarea de "tendință către mediocritate" dată acestor observații suscită încă şi azi o serie de discuții.

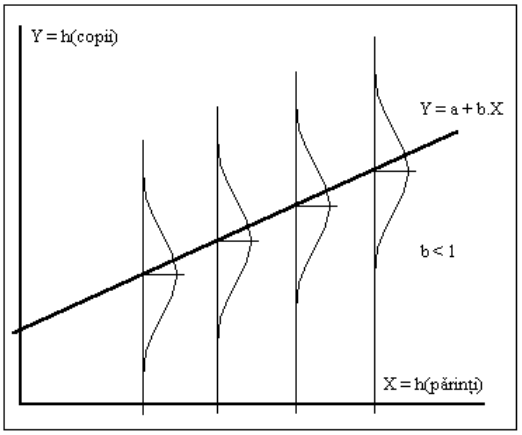


Figura II.24. Alura unei reprezentări ilustrând termenul de "regresie"

viii - Testarea liniarității

Corelația liniară este cea mai simplă şi cea mai studiată, de aceea în analiza corelației ea se efectuează prima; deseori, obținerea unor rezultate ce indică o corelație liniară slabă este interpretată - nejustificat - ca absență a unei corelații. Acest lucru poate fi adevărat însă există dese situații când variabilele sunt destul de puternic corelate însă nu liniar (figura II.25.).

Pentru a verifica liniaritatea se construieşte o nouă variabilă:

$$Z_i = \frac{Y_i - \hat{Y}_i}{\sqrt{S^2}} = \frac{\varepsilon_i}{\sqrt{S^2}} \qquad \text{(II.6.14.a)}$$

. Ipoteza de zero: H_0: regresia este liniară

. Test aplicat: testul Z al distribuției normale, astfel:

- alegem un nivel de semnificație α şi luăm din tabel Z_α

- dacă $|Z_i| < Z_\alpha \ \forall i = 1...N$ (II.6.14.b)

atunci acceptăm H_0, în caz contrar o respingem.

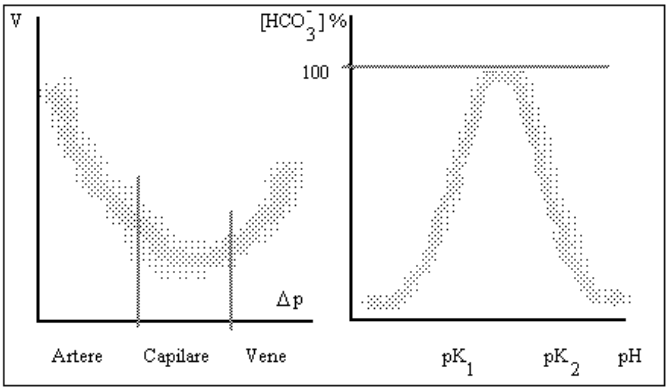


*Figura II.25. Co*relația neliniară

ix ^{0} **Încadrarea dreptei de regresie**

Dreapta de regresie teoretică $Y = \alpha + \beta X$ poate lua valori în intervalul (aici α = ordonata la origine, β = panta)

$$Y \in (\hat{Y}\text{-}t \cdot S_{\hat{y}}, \hat{Y} + t \cdot S_{\hat{y}}) \qquad \text{(II.6.15)}$$

unde t este valoarea din tabelul repartiției t pentru un prag de semnificație ales.

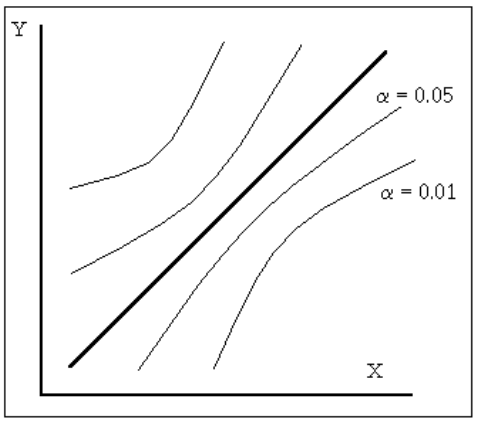


Figura II.26. Încadrarea dreptei de regresie în intevale de încredere de 95% şi 99%

Eroarea standard a estimării lui $\hat{Y}$ pentru fiecare X se calculează cu:

$$S_{\hat{y}} = \sqrt{S^2} \cdot \sqrt{\frac{1}{N} + \frac{\sum (X - \bar{X})^2}{\sum X^2 - \frac{(\sum X)^2}{N}}} \qquad \text{(II.6.16)}$$

În felul acesta în reprezentarea grafică a diagramei se trasează şi limitele de încadrare a dreptei (figura II.26.).

Observație: Dreapta de regresie a variabilei *Y* față de *X* este diferită de dreapta de regresie *X* față de *Y* (deci cea care s-ar obține dacă am inversa axele pe care sunt plasate cele două variabile). De aceea, când ar exista pericol de confuzie, coeficienții dreptei $y = f(x) = a + bx$ se notează $b_{y/x}$ şi $a_{y/x}$ în timp ce pentru dreapta $X = f(y)$ se notează $b_{x/y}$ respectiv $a_{x/y}$. Coeficientul de corelație **r** rămâne acelaşi în ambele situații.

C. Corelații si regresii neliniare

Deşi corelația liniară este întâlnită destul de des, o serie de fenomene din meteria vie conduc la reprezentări destul de depărtate de o dreaptă, astfel încât este mult mai potrivită alegerea altei relații pentru descrierea dependendenței între variabile în cazurile respective.

a. Raport de corelație

În cazul regeresiei neliniare, în locul coeficientului de corelație **r** se foloseşte un alt parametru numit raport de corelație, dat de formula:

$$\eta_{xy} = \sqrt{1 - \frac{\sum \varepsilon_i^2}{\sum \tau_i^2}} = \sqrt{\frac{\sum \delta_i^2}{\sum \tau_i^2}} \qquad \text{(II.6.17)}$$

unde ε , δ şi τ au semnificaţia din figura II.22., cu deosebirea că punctul P' va fi situat pe curba de regresie (deci nu neaparat pe o "dreaptă"). În această relaţie $\hat{y}$ se va calcula conform curbei care se presupune că descrie relaţia dintre *x şi y.* Dacă avem o corelaţie liniară, raportul de corelaţie va fi egal cu coeficientul de corelaţie.

Intervalele de încredere pentru rapoartele de regresie se calculează cu ajutorul coeficienţilor F din testul lui Fisher.

În cele ce urmează vom enumera câteva corelaţii neliniare mai des întâlnite în medicină şi biologie.

b. Corelaţii si regresii exponenţiale

Sunt foarte des întâlnite în descrierea fenomenelor naturale.

- Ecuaţia regresiei exponenţiale:

$$y = a \cdot e^{bx} \qquad \text{(II.6.18)}$$

având coeficienţii *a şi b.*

- Exemple:

i ^{0} - corelaţii exponenţiale crescătoare (b > 0)
- în fenomene de absobţie (intenstinală etc.)

ii - corelaţii exponenţiale descrescătoare (b < 0)
- clearance - funcţia de epurare (renală, hepatică)

c. Corelaţii şi regresii logaritmice

- Ecuaţia regresiei logaritmice:

$$y = a + b \cdot \log x \qquad \text{(II.6.19)}$$

- Exemple:

. legea Weber - Fechner - între senzaţie şi intensitatea stimulului.

d. Corelaţii şi regresii ca funcţie putere

- Ecuaţia funcţiei putere:

$$y = a \,.\, x^{b} \qquad \text{(II.6.20)}$$

- *Exemple:*

. legea lui Stevens - între frecvenţa impulsurilor nervoase pe o fibră şi intensitatea stimulului

e. Corelaţii si regresii hiperbolice

- Ecuaţia funcţiei hiperbolice:

$$(x+a)\,(y+b) = k \qquad \text{(II.6.21)}$$

- *Exemple:*

. legea lui Hill - relaţia între forţă şi viteza de contracţie pentru muşchiul striat

. legea lui Abbey - relaţia între intensitatea şi durata unui stimul luminos foarte scurt pentru determinarea pragului de sensibilitate.

f. Corelaţii şi regresii logistice

- Ecuaţia funcţiei logistice:

$$y = \frac{a \cdot x}{b + x} \qquad \text{(II.6.22)}$$

- *Exemple:*

. Cinetica Michaelis - Menten - relația între viteza reacției enzimatice şi concentrația de substrat.

. Curbele doză-efect-relația între doza unei substanțe medicamentoase şi efectul dozei respective asupra unui țesut (Ariens)

(Observație: reprezentările funcției logistice se fac de obicei în coordonate y = f (log x), funcția având în acest caz o formă sigmoidală şi o serie de proprietăți de simetrie).

Există şi alte tipuri de regresii cu care ne mai putem întâlni: parabolice, polinomiale etc.

D. Metode de fitare

Definiție: Metodele folosite pentru a găsi "cea mai bună" dreaptă, sau curbă de un anumit tip, care să treacă printre punctele experimentale se numesc metode de fitare.

Cele mai des întâlnite metode de determinare a parametrilor curbei (dreptei) de regresie sunt:

a. Metoda celor mai mici pătrate, pe care am descris-o anterior, bazată pe minimizarea sumei abaterilor punctelor experimentale de la curba de regresie (formula 6.8.a.)

b. Metoda transformărilor liniare, prin care se efectuează în ecuația curbei de regresie o schimbare de variabilă astfel încât, cu noile variabile reprezentarea să devină o dreaptă. Iată câteva exemple:

- pentru regresia exponențială

$\log y = z, \log a = c : \; z = c + b \cdot x$ (II.6.18)"

- pentru regresia logaritmică

$\log x = z : \; y = a + b \cdot z$ (II.6.19)"

- pentru regresia putere

$\log y = z , \log x = t, \; \log a = c : \; t = c + b \cdot t$ (II.6.20)"

- pentru regresia hiperbolică

$\frac{1}{x+a} = z : \; y = -b + k \cdot z$ (II.6.21)"

- pentru corelația logistică

$\frac{1}{y} = z, \quad \frac{1}{x} = t, \quad \frac{1}{a} = c, \quad \frac{b}{a} = d : \; z = c + d \cdot t$ (II.6.22)"

(Observație: această transformare liniară se mai numeşte transformarea Lineweaver Burke sau transformare dublu reciprocă şi este mult utilizată în prelucrarea datelor de cinetică enzimatică).

Trebuie menționat că metoda transformărilor liniare conduce la rezultate ce diferă de metoda celor mai mici pătrate aplicată direct la datele experimentale.

c. Metoda asemănării maxime - se bazează pe determinarea valorilor pentru care datele experimentale ar fi apărut aşa cu cea mai mare probabilitate. Rezultatele obținute sunt apropiate de cele din metoda celor mai mici pătrate.

6.2. RELAȚII ÎNTRE DOUĂ VARIABILE ORDINALE

În cazul variabilelor ordinale parametrii definiți anterior nu se mai potrivesc şi sunt definite mărimi specifice pentru ranguri.

A. Coeficientul de corelație a rangurilor - Spearman

Este un coeficient de corelație liniară între rangurile acordate diferiților "indivizi" în clasificări diferite.

a. Formula

$$R = 1 - \frac{\sum D_{:i}^2}{N(N^2-1)} \qquad \text{(II.6.23)}$$

unde D_i este diferența între rangurile individului *i* în cele două clasificări.

b. Exemplul II.21.a

Considerăm rezultatele obținute prin testul psihologic Luscher de preferință a culorilor pe două loturi: un grup de adulți și un grup de copii (cu vârsta 5-15 ani) - tabelul II.19.

Tabel II.19.a. Rangurile preferinței culorilor prin testul Luscher la două loturi

Culoarea	***Rangul preferinței***		***D***	***D*** ^{2}	***Rezultate:***
	Copii	***Adulți***			
R=Roşu	1	5,5	-4,5	20,25	N=6
G=Galben	2	5,5	-3,5	12,25	$\sum D^2 = 51$
V=Verde	5	4	1	1,00	R=0,35
A=Albastru	3,5	1	2,5	6,25	$R_{0,05}^{tab} = 0{,}829$
W=Alb	3,5	2	-1,5	2,25	$R_{0,01}^{tab} = 0{,}943$
N=Negru	6	3	3	9,00	*==>Corelație nesemnificativă*

e. Semnificația coeficientului de corelație Spearman

La fel ca şi în cadrul coeficientului de corelație liniară, coeficientul R poate fi comparat cu valori dintr-un tabel.

- Ipoteza de zero: H_0 : R = 0
- Test aplicat: test specific pentru R (coeficientul de corelație al ragurilor).
- Aplicație: pentru un nivel de semnificație α *(0,05 sau 0,01)* se caută valoarea din tabel $R_{\alpha,N} = R_{tab}$.

Dacă $|R| < R_{tab}$ se acceptă $|H_0|$ şi se consideră corelația nesemnificativă, în caz contrar se respinge H_0 şi se consideră o corelație semnificativă a rangurilor.

B. Coeficientul de corelație Kendall

Este tot un coeficient de corelație pentru ranguri.

a. Formula

$$K = \frac{2S}{N(N-1)} \qquad \text{(II.6.24)}$$

unde S este suma scorurilor pozitive şi negative ale rangurilor dintr-o clasificare în raport cu cealaltă clasificare.

b. Exemplul II.21.b:

Rearanjăm datele din tabelul II.19.a. astfel încât o clasificare să fie ordonată (de ex. cea pentru copii); datele apar acum ca în tabelul II.19.b.

Deci $S = -9 + 5 = -4$ şi înlocuind în (II.6.24) obţinem K = - 0,266.
Există tabele prin care se poate în continuare verifica şi semnificaţia acestui coeficient de corelaţie.

Tabel II.19.b Rangurile preferinţei culorilor - două clasamente obţinute pe două loturi: copii şi adulţi - rearanjarea datelor din tabelul 7.1

Culoarea	*Rangul preferinţei*		*Notaţie*	*D(-)*	*D(+)*
	Copii	*Adulţi*		$r_i \langle r_{j,}, j > i$	$r_i \rangle r_j, j \rangle i$
Roşu	1	5,5	r_1	**4**	**0**
Galben	2	5,5	r_2	$(r_3, r_4, r_5, r)_6$	**0**
.....	.....	.....	.....		.....
Albastru	3,5	1	r_3	**4(r**	**3**(r_4, r_5, r_6)
Alb	3,5	2	r_4	$_3, r_4, r_5, r_6)$	**2**(r_5, r_6)
.....	.....	.....	.....	.....	.....
Verde	5	4	r_5	**0**	**0**
Negru	6	3	r_6	**0**	-
				.....	
				1(r_6)	
				-	
				-9	**+5**

6.3. RELAŢII ÎNTRE VARIABILE NOMINALE

Corelaţia între variabilele nominale nu se caracterizează prin coeficienţi de corelaţie ci se efectuează prin aplicarea unui test statistic care să indice prezenţa / absenţa unor corelaţii între clasificările realizate după mai multe criterii.

A. Testul de independenţă χ^2

Sub forma prezentată în capitolul 5, testul χ^2 poate da informaţii asupra corelaţiei / independenţei între clasificările în câte două clase, după două criterii.

Testul poate fi generalizat pentru *m* clase şi *n* criterii; în acest caz se utilizează mai des un **coeficient de contingenţă C:**

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}} \qquad \text{(II.6.25)}$$

Coeficientul C = 0 indică independenţa variabilelor. Cu cât este mai mare, cu atât legătura este mai puternică (valoarea maximă $C_{\max} = 1/\sqrt{2} = 0{,}707$).

B. Alți indicatori

Pentru variabile nominale sau propus şi alți indicatori care să ilustreze posibile relații între clase:

a. Indicatori de asociere

- folosit pentru tabele de contingeță 2 * 2

- formula:

$$\varphi = \frac{bc - ad}{\sqrt{L_1 L_2 C_1 C_2}}$$, L_1, L_2, C_1, C_2 fiind totalurile pe linii, respectiv coloane

- $\varphi \in [-1, +1]$; valori extreme indică asociere puternică, valori în jurul lui 0 indică independență

- semnificația statistică se determină cu ajutorul repartiției χ^2, statistica fiind calculată cu formula:

$$\chi^2 = N \cdot \varphi^2 \qquad \text{(II.6.26.b)}$$

b. Indicatori de grupare

Prin diverse tipuri de analize se pot găsi criterii după care indivizii unui lot se pot grupa în mai multe clase astfel încât să se poată preciza asemănarea între indivizii unei clase şi deosebirea lor față de indivizii altor clase.

6.4. RELAȚII ÎNTRE MAI MULTE VARIABILE CANTITATIVE

În cazul în care generalizăm analiza bivariată, în care urmăream relația între variabilă (dependentă) şi o variabilă independentă, obținem o **analiză multivariată,** în care avem o funcție de mai multe variabile:

$$y = f(x_1, x_2, \ldots, x_n) \qquad \text{(II.6.27)}$$

Cea mai simplă relație este **regresia liniară multiplă**, în care considerăm dependența de forma:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n \qquad \text{(II.6.28)}$$

Cea mai bună suprafață de regresie se obține când:

$$\sum \varepsilon_i^2 = \sum (y^i - \hat{y})^2 = min \qquad \text{(II.6.29)}$$

unde $\hat{y}^i = b_0 + b_1 x_1^i + b_2 x_2^i + \ldots + b_n x_n^i$ (II.6.28)'

Pentru **regresia multiplă** se definesc:

- coeficientul de corelație global
- coeficienții de corelație parțiali (luând pe rând fiecare pereche de variabile).

7. EPIDEMIOLOGIE

Epidemiologia este un domeniu medical pluridisciplinar având o zonă de intersecție mare cu biostatistica. În epidemiologia clinică se urmăreşte atât determinarea frecvenței de apariție a unei boli cât şi definirea unor asocieri între boală şi factori cauzali sau favorizanți. Când se suspectează vreo astfel de asociere, se încearcă la început să se identifice condițiile care determină creşterea riscului unei afecțiuni, apoi evidențierea unei relații cauză-efect, având în final consecințe în dezvoltarea unui tratament adecvat şi a unor strategii profilactice.

Studiile epidemiologice intră în categoria studiilor populaționale care cuprind două mari capitole:

- analiza riscului (partea centrală a epidemiologiei)
- analiza supraviețuirii

7.1. ANALIZA RISCULUI

A. Factori de risc

a. Definiție: O cauză ipotetică (indiferent de natură - comportament, condiție, caracteristică fizică sau de mediu etc.) ce determină creşterea probabilității ca un individ sănătos să dezvolte a anumită boală reprezintă un factor de risc.

b. Clasificare:

- factori de mediu: factori poluanți, toxine, microorganisme infecțioase etc.
- factori comportamentali (obiceiuri): fumat, alcool, droguri, nerespectarea măsurilor de protecție a muncii, sedentarism etc.
- factori sociali: evenimente familiare tragice, divorț, pierderea serviciului etc.
- factori genetici: hipercolesterolemie etc.

c. Tipuri de expunere la acțiunea factorului de risc:

- expunere punctuală - ex. accidente (la o intreprindere chimică etc.)
- expunere cronică - cea mai frecventă; se estimează în aceste condiții "doza" curentă, doza cumulată, durata expunerii etc.

d. Relația factor risc / boală:

- factor cauzal - când putem atribui factorului o acțiune directă
- factor favorizant (marker) care creşte probabilitatea, dar nu i se poate atribui o acțiune directă (ex: factorii sociali - economici, educaționali etc.).

B. Prezentarea datelor

Uzual datele din analiza riscului se prezintă sub forma unui tabel de contingență, cel mai frecvent 2 x 2 (cu două linii şi două coloane) în care întregul lot de N indivizi este împărțit în grupul de indivizi expuşi (L1), respectiv neexpuşi (L2). Din fiecare grup, o parte dezvoltă boala, (N11 din L1, respectiv N21 din L2), o parte nu (N12 din L1, respectiv N22 din L2) - tabelul II.20

Tabel II.20. Prezentarea schematică a datelor unui studiu epidemiologic: E+ =expuşi, E- = neexpuşi la acțiunea factorului de risc; B+ = prezintă boala, B- = nu prezintă boala

		B+	B -	
Expunere	E+	N11	N12	L1
	E -	N21	N22	L2
		C1	C2	N

C. Metode de studiu în epidemiologie

a. Studii experimentale

Din punct de vedere teoretic rezultatele cele mai de încredere s-ar obține într-un studiu experimental, în care investigatorul are controlul complet asupra factorului de risc (ca variabilă independentă, cu rol cauzal) şi urmăreşte efectul asupra grupelor (varibile dependente). Din considerente etice însă, aceste studii sunt limitate doar la acțiunea unor factori cu risc redus.

b. Studii observaționale

În marea majoritate a cazurilor studiile trebuie efectuate pe loturi în care expunerea nu s-a întâmplat la dorința expresă a investigatorului. Marile dezavantaje ale studiilor observaționale sunt: precizia limitată a măsurării acțiunii factorului de risc (intensitate, durată) şi stabilirea grupelor expuşi / neexpuşi ce vor fi comparate concret. Recunoaşterea şi "controlul" unor eventuale surse de "bias" constituie unul din elementele urmărite prioritar în aceste analize.

Dintre tipurile de studii observaționale mai des întâlnite enumerăm:

i - Studiul transversal ("cross-secțional") se mai numeşte şi studiu de prevalență. Este cel mai simplu model, bazat pe fotografierea unei situații la un moment dat, culegând date de tipul celor din tabelul II.20.

Dintre dezavantajele mai des citate reținem:

- estimarea prevalenței este influențată în cazul evoluțiilor rapide (fie spre deces fie spre recuperare)
- incertitudinea "antecedență - consecință"

ii - Studiul prospectiv pe cohortă ("cohort prospective"/"folow-up"/ "longitudinal"). Pornind de la numele unei unități militare în epoca romană (cohorta), care oferea condiții asemănătoare de luptă pentru membrii ei, studiile de acest tip iau în analiză două loturi din persoane inițial sănătoase, dintre care unul este supus la acțiunea factorului de risc suspectat. Loturile sunt urmărite în timp, pornind din momentul definirii lor; se identifică apariția afecțiunii în ambele loturi. (figura II.27.a)

iii - Studiul retrospectiv pe cohortă ("historical cohort / retrospective / non-concurrent"). La fel ca în studiul prospectiv pe cohortă evoluția se urmăreşte în sensul natural al scurgerii timpului, pornind de la situația unui grup inițial din care o parte au fost expuşi şi acum putem evalua la câți din fiecare grup, a apărut afecțiunea analizată (figura II.27.a)

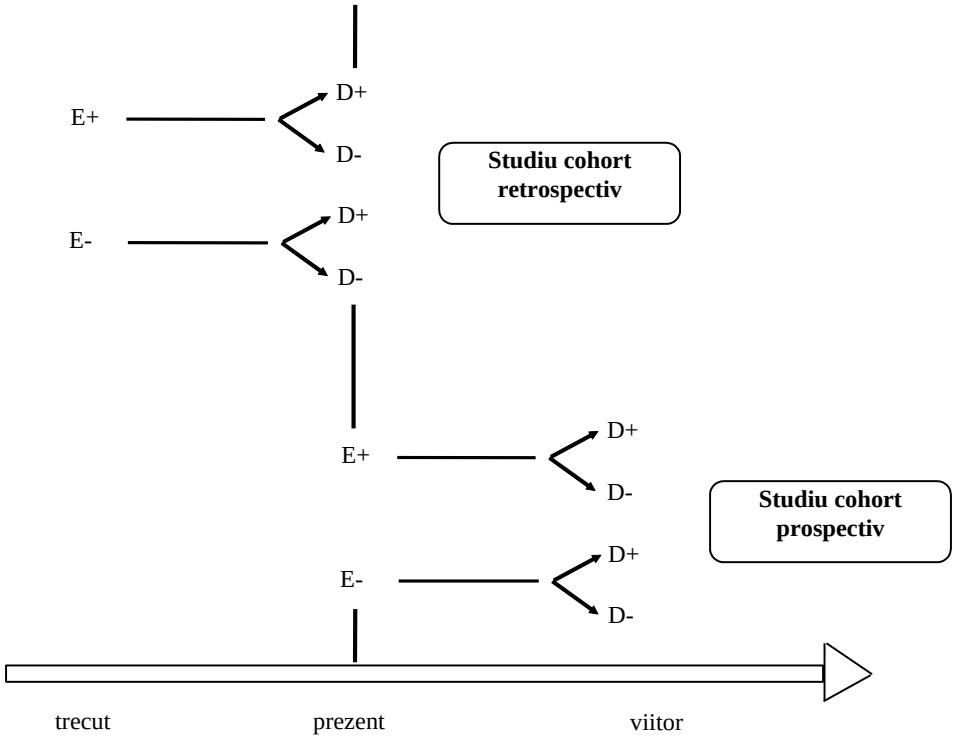
iv - Studiul retrospectiv clasic ("case-control") - în care grupul analizat cuprinde "cazurile" în care a apărut boala şi investigăm în care din aceste cazuri a existat o expunere la factorul de risc (deci urmărim în sens invers temporal) şi identificăm apoi prezența/absența factorului de risc şi pentru un grup martor (control) - figura II.27.b. Aceste studii "case-control", deşi mai comode pentru colectarea unor date, au destule dezavantaje: grupul martor nu poate fi întâmplător ci trebuie selectat pe aceleaşi criterii ca şi grupul de "cazuri"; de asemenea, dacă grupul B+ este selectat dintre cazurile spitalizate, el deja cuprinde un important *bias*: cazurile mai grave!

c. Compararea metodelor

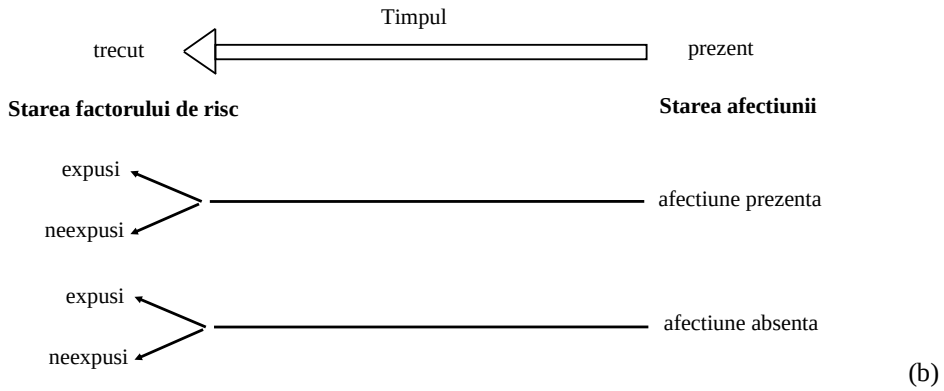
O analiză a posibilelor surse de erori în diversele tipuri de studii ne permite o ierarhizare a metodelor enumerate, cele mai bune rezultate fiind aşteptate de la studiile experimentale; prezentarea ierarhică a metodelor este schițată în tabelul II.21.

Tabel II.21. Ierarhia metodelor de studiu în epidemiologie, functie de puterea acestora.

experimental > cohort-prospectiv > cohort-retrospectiv > case-control > cross-sectional



(a)



(b)

Figura II.27. Tipurile de studii epidemiologice - prezentarea schematică: (a) studii de tip cohort; (b) studii de tip *case-control*

D. Parametri fundamentali în epidemiologie

a. Indici în studii populaționale

i - Prevalența unei boli într-o populație: este proporția din populația respectivă având boala (la un moment dat):

$$\Pr v\,(B,t) = \frac{N_{B\,+(t)}}{N} = \frac{Nr.indivizi\ \ având\ \ boala\,B}{nr.\ populatie} \qquad \text{(II.7.1)}$$

ii - Incidența unei boli: este numărul de cazuri ce apar într-un interval Δt (t_1 ,t_2) într-o populație cu risc. Ea poate fi exprimată prin:

. incidența cumulativă CI: proporția într-un grup fix predefinit (cohortă) la care apare boala în intervalul specificat (fig. II.28).

$$\text{CI (B, } \Delta t) = \frac{N^{inc}(\Delta t)}{N_{risc}} = \frac{\text{nr. cazuri noi in } \Delta\,\text{t}}{\text{nr. populatie cu risc}} \qquad \text{(II.7.2)}$$

Pentru exemplul din figura II.28.

$$Prv\ (1\ ian.95)\ =\ \frac{4}{100} = 4\%$$

$$CI\ = \frac{7\ (cazuri\ noi)}{100 - 4\ (cazuri\ existente\ la\ momentul\ t_1)} = \frac{7}{96} = 7{,}4\%\ pe\ perioada\ de\ 1\ an$$

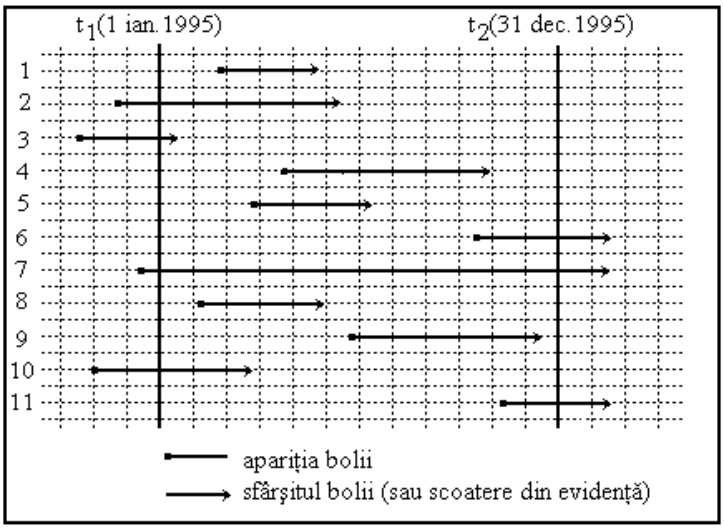


Figura II.28. Cazuri de hepatită B într-un lot de 100 persoane: calculul indicelui cumulative

. densitatea de incidență ID (numită şi rata de incidență, rata de hazard sau "forța" morbidității/mortalității): este dată de numărul de cazuri noi ce apar într-un interval Δt (t_1,t_2) într-o populație cu risc, studiată pe diverse perioade de timp. Într-un studiu practic, pe o perioadă îndelungată, din lotul inițial (tip cohortă) se pierd o serie de persoane din diverse motive (se mută, mor din alte motive , nu continuă tratamentul

etc.). De aceea, cei care nu au fost prezenți întreaga perioadă nu se scot din studiu ci vor fi luați în considerare numai în măsura în care situația lor a fost cunoscută.

$$ID\ (B\ ,\Delta t) = \frac{N^{inc}(\Delta t)}{N^{*}_{risc}} = \frac{\text{nr. cazuri noi in } \Delta t}{\text{nr. mediu populatie cu risc pe interval}} \qquad \text{(II.7.3)}$$

În figura II.29 este prezentat un model de evidență extins față de cel din figura II.28.

Din cohorta de 15 persoane cu risc (4 la 1 ian.1992, apoi 5 noi cazuri, 2 şi respectiv 4 noi cazuri la fiecare început de an 93-95), la 9 a fost depistată afecțiunea (1,3,6,7,8,9,10,12,14), care a determinat decesul în 5 cazuri (1,7,9,12,14); 2 au supraviețuit până la încheierea studiului (3,6), 1 a fost pierdut din evidență (8), iar unul a decedat din alte cauze (10). Din restul de 6 persoane cu risc studiate (2,4,5,11,13,15) la încheierea celor 5 ani de urmărire mai erau în evidență 4 (5,11,13,15), unul a fost pierdut din evidență (2) şi unul a decedat din alte motive (4). Numărul total de ani de risc pe întregul lot este suma coloanei din dreapta: 35,5 ani. Deci:

$$ID = \frac{9}{35.5} = 0.25\ \text{cazuri/(persoana * an)}$$

Dacă loturile sunt omogene se poate folosi cu aproximație relația:

$CI \approx ID\ x\ \Delta t$ (II.7.4)

De asemenea, se poate aproxima o relația între prevalență şi incidență:

$Prr \approx ID\ x\ \Delta T\ (B)$ (II.7.5)

unde ΔT (B) este durata medie a bolii.

iii Rata de morbiditate (Mrb): este incidența unei boli într-o populație, într-un anumit interval de timp (adesea 1 an).

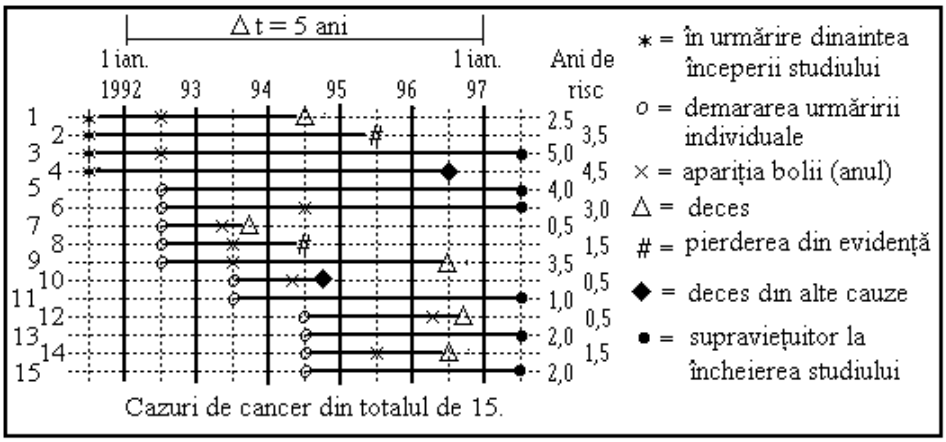


Figura II.29. Studiu de urmărire a evoluției unui lot cu risc de cancer

iv Rata de mortalitate: enumeră cazurile de deces dintr-o populație, într-un anumit interval de timp; se utilizează:

- rata de mortalitate generală: din orice cauză
- rata de mortalitate specifică pe cauze: separat, pe boli sau grup de boli - (de exemplu: cardiovasculare etc.)
- rata brută de mortalitate: față de întreaga populație
- rata de mortalitate specifică pe categorii: separat, pe anumite subgrupe de populație
- rata de mortalitate pe grupe de vârstă

- rata de mortalitate corectată, pe grupe de vârstă - se fac corecţii în funcţie de distribuţia pe grupe de vârstă.

v Rata de fatalitate a bolii: rata de deces într-o populaţie având boala, într-un interval de timp.

vi Rata de atac: pentru boli cu durată scurtă, când durata observaţiei acoperă întreaga "epidemie", proporţia celor ce dezvoltă boala din populaţia cu risc (= CI).

b. Parametrii în analiza riscului

i - Indicele "odd" (suportul "succes / eşec"): este probabilitaqtea de apariţie a bolii la cei expuşi faţă de probabilitatea de a nu apărea boala la cei expuşi. Cu notaţiile din tabelul II.21 putem scrie:

$$ODD\ (E+) = N11 / N12 \qquad \text{(II.7.6)}$$

$$ODD\ (E\ -) = N21 / N22$$

adică în câte cazuri prezenţa factorului de risc (E+ = subiecţi expuşi) are "succes" în declanşarea bolii (N11) faţă de situaţiile de "eşec" (N12); similar raportul succes / eşec pentru condiţia absenţei factorului de risc (E - = neexpuşi).

ii - Raportul odds ("odds ratio"): este raportul indicelui "odd" pentru grupul expus faţă de cel neexpus la factorul de risc:

$$OR = \frac{ODD(E+)}{ODD(E-)} = \frac{N11/N12}{N21/N22} = \frac{N11 \cdot N22}{N21 \cdot N12} \qquad \text{(II.7.7)}$$

iii - Riscul relativ: este probabilitatea de apariţie a afecţiunii la cei expuşi faţă de probabilitatea de apariţie a afecţiunii la la cei neexpuşi la factorul de risc.

$$RR = \frac{N11/L1}{N21/L2} \qquad \text{(II.7.8)}$$

Dacă riscul relativ are valoarea RR ≈ 1 putem spune că factorul analizat nu reprezintă un factor de risc, probabilitatea de apariţie a afecţiunii fiind la fel de mare şi la lotul neexpus factorului de risc. Valori RR >1 dau o semnificaţie acţiunii factorului de risc. Pentru a estima intervalul de încredere în care parametrul RR poate fluctua întâmplător se folosesc "limitele Cornfield" pentru p = 95% probabilitate ca ipoteza de zero să fie adevărată.

iv - Riscul atribuabil: este diferenţa între probabilitatea de apariţie a bolii la cei expuşi şi cea de apariţie a bolii la cei neexpuşi. Formula este:

$$AR = N11 / L1 - N21 / L2 \qquad \text{(II.7.9)}$$

În funcţia de tipul studiu efectuat se mai pot defini şi alţi parametri pe care nu-i mai prezentăm aici.

E. Analiza multistratificată

Deseori indivizii unei populaţii sunt supuşi la acţiunea simultană a mai multor factori de risc. Depistarea contribuţiei fiecărui factor de risc la efectul final se realizează prin analiza multistratificată. Se alcătuiesc tabele de forma celui din tabelul II.22.a.

Tabel II.22.a. Prezentarea datelor într-un studiu cu doi factori de risc (fumat, cafea)

Lot: boală coronariană +			Lot: boală coronariană -		
Subiect	Fumat	Cafea (mg / zi)	Subiect	Fumat	Cafea - (mg / zi)
1	DA	1100	1	DA	1000
2	DA	800	2	NU	300
3	NU	200	3	NU	100
.....	.....	.....	.....	.....	.....
12		8/4	12	3/9	$\overline{X} = 400$
$\overline{X} = 716$					

O analiză superficială incompletă ne-ar putea induce ideea unui risc crescut al consumului de cafea asupra declansării afecţiunilor coronariene, conform centralizării din tabelul II.22.b.

Tabel II.22.b. Influenţa consumului de cafea în bolile coronariene

Boală	Consum zilnic mediu (mg)
B +	716
B -	400

O stratificare încât să se includă şi fumatul, va scoate în evidenţă rolul dominant al acestuia (tabel II.22.c.).

Tabel II.22.c. Tabel stratificat: consumul mediu de cafea / zi la fumători şi nefumători, respectiv coronarieni şi necoronarieni

Fumat	B +	B -	Medie
DA	950 (n = 8)	1000 (n = 3)	963 (n = 11)
NU	250 (n = 9)	200 (n = 9)	216 (n = 13)
Medie	716 (n = 12)	400 (n = 12)	558 (n = 24)

Analizele multistratificate sunt destul de dificile; uneori este greu a discerne între factorul cauzal şi alţi factori asociaţi. Există nişte criterii definite de Hill care ar facilita această operaţiune.

7.2. ANALIZA SUPRAVIEŢUIRII

Un succes indiscutabil al medicinei moderne îl prezintă rezultatele tratamentelor în cazurile cu diagnostice severe. Depistarea precoce a afecţiunilor grave şi lărgirea paletei şi eficienţei tratamentelor au generart extinderea sensibilă a speranţei de viaţă după diagnosticarea bolii. Estimarea eficienţei unor terapii şi compararea tratamentelor se realizează prin studii epidemiologice. Deşi pot fi retrospective, majoritatea studiilor sunt în general prospective, o serie de date necesare pentru analiza statistică nefiind disponibile pentru studiile retrospective. Capitolul din epidemiologie referitor la aceste studii, numit "analiza supravieţuirii" şi-a extins sfera de aplicabilitate şi asupra altor tipuri de studii în care se urmăreşte pe o durată mare de timp (luni, ani) rezulatatul unei terapii.

Studiile de acest gen au fost iniţial solicitate de companiile de asigurări, ulterior devenind un capitol bine definit în epidemiologie.

Metodologia acestor studii a fost standardizată, OMS publicând în 1974 recomandările UICC ("Union Internationale Contre le Cancer"): regulile TNM (tumori, noduli, metastaze).

A. Caracteristicile studiilor de lungă durată

Studiile recomandate sunt de tip "cohort prospectiv", însă în cazul unor perioade îndelungate (5-20 ani) apar o serie de factori de care trebuie să ţinem seama:

- o serie de indivizi din lotul iniţial pot fi pierduţi din evidenţă (îşi mută domiciliul, intervin alte tratamente etc.); aceste "date lipsă" pot să reprezinte uneori un procent însemnat din ansamblul de date; pentru prelucrări aceste cazuri nu se abandonează ci se iau în considerare, dar numai pentru intervalul de timp pentru care situaţia individului este clar cunoscută;
- persoanele din lot trăiesc în condiţii diferite astfel încât această heterogenitate face mai greu vizibil efectul datorat numai factorului de risc;
- foarte des din ansamblul condiţiilor putem desprinde unele care pot fi de asemenea considerate "factor de risc", ce acţionează sinergic sau competitiv cu factorul urmărit de noi.

B. Prezentarea şi prelucrarea datelor

a. Tabele de viaţă

Metodologia OMS sugerează colectarea datelor pentru prelucrare sub forma unor "tabele de viaţă" ("life tables"):

Exemplu II.22. Datele sunt redate în tabelul II.23. Iată descrierea coloanelor:

1. Anul de observaţie ($i \div i+1$): se calculează numărul de ani de la data începerii tratamentului; de ex: un pacient care a fost prima dată tratat în 7 aprilie 1947 şi a decedat în 24 noiembrie 1950 va fi considerat decedat în intervalul 3-4.
2. În viaţă la începutul intervalului (l_i): primul număr (1000) indică numărul total de pacienţi studiaţi; nu înseamnă că toţi au început tratamentul în aceeaşi zi; ei sunt luaţi în evidenţă pe măsură ce sunt depistaţi şi încep tratamentul; în tabel intervalele se măsoară pentru oricare pacient pornind de la ziua primului tratament. Numărul de indivizi cunoscuţi a fi în viaţă la începutul fiecărui interval (început de nou an de la luarea în evidenţă) se calculează din precedentul scăzând d_i, u_i şi w_i , deci:

$$l_{i+1} = l_i - (d_i + u_i + w_i) \qquad \text{(II.7.9.a)}$$

3. Cei decedaţi *(d_i)* datorită bolii în intervalul $i \div i + 1$.
4. Pierduţi din urmărire *(u_i)*: aici se includ cei a căror situaţie, la data încheierii studiului (31 decembrie 1960) nu este cunoscută, însă pentru care este cunoscută starea până la un moment dat; de exemplu: un pacient care a început tratamentul în 20 Septembrie 1946 şi era în viaţă pe 6 iunie 1949, după care nu se mai ştie nimic, va fi considerat pierdut în intervalul 2-3. Aici sunt de obicei incluşi şi cei decedaţi din alte cauze.
5. Scoşi din urmărire fiind în viaţă la sfârşitul perioadei analizate *(w_i)*. În exemplul nostru perioada analizată se încheia la 31 decembrie 1960; un pacient care a început tratamentul în 5 mai 1954 şi este în viaţă la 31 decembrie 1960 va fi scos din calcul în intervalul 6-7 (a supravieţuit 6 ani şi n-a fost urmărit mai mult).
6. Numărul efectiv al celor expuşi la riscul de deces *(n_i)*. Pacienţii pierduţi din urmărire *(u_i)* şi cei scoşi din urmărire *(w_i)* sunt consideraţi ca fiind distribuiţi uniform de-a lungul întregului an, ei pot fi consideraţi ca expuşi timp de jumătate de interval; deci:

$$n_i = l_i - (\mu_i + w_i)/2 \qquad \text{(II.7.9.b)}$$

S-a presupus că probabilitatea de supravieţuire pentru cei pierduţi sau scoşi din urmărire este aceeaşi ca şi pentru cei rămaşi în evidenţă.

7. Rata anuală de mortalitate *(q_i)* reprezintă proporţia celor decedaţi în fiecare an calculată ca probabilitate de deces:

$$q_i = d_i / n_i \qquad \text{(II.7.9.c)}$$

8. Rata anuală de supravieţuire *(p_i)* reprezintă probabilitatea de a supravieţui în intervalul $i \div i+1$ (calculată pentru cei în viaţă la începutul intervalului):

$$p_i = 1 - q_i \qquad \text{(II.7.9.d)}$$

9. Rata cumulativă de supravieţuire de la început până la inclusiv intervalul $i \div i + 1$ se calculează cu:

$$p_i = p_1 \cdot p_2 \cdot \ldots\ldots \cdot p_i = \prod_{j=1}^{i} p_j \qquad \text{(II.7.9.e)}$$

Tabelul II.23. Prezentarea datelor pentru prelucrarea prin metoda actuarială sub formă de "tabele de viaţă". Exemplul se referă la un studiu la pacienţi care au început tratamentul între 1946-1955 şi urmăriţi până la 31 decembrie 1960. (* din [UICC - TNM])

1	Anul de observaţie	i - i+i	0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11
2	În viaţă la începutul intervalului	l_i	1000	816	642	558	504	460	382	308	259	212	165
3	Decedaţi în interval	d_i	180	170	80	50	40	28	26	7	7	11	
4	Pierduţi din urmărire în interval	u_i	4	4	4	4	4	6	5	4	3	3	
5	În viaţă la sfârşitul intervalului şi scoşi din urmărire	w_i	-	-	-	-	-	44	43	38	37	33	165
6	Numărul efectiv de expuşi la riscul de deces	n_i	998	814	640	536	502	435	358	287	239	194	
7	Rata anuală de mortalitate	q_i	0.180	0.209	0.125	0.090	0.010	0.064	0.073	0.024	0.029	0.057	
8	Rata anuală de supravieţuire	$p_i = 1-q_i$	0.820	0.791	0.875	0.910	0.920	0.936	0.927	0.976	0.971	0.943	
9	Rata cumulativă de supravieţuire până la sfârşitul intervalului	$P_i = p_1 \cdot p_2 \cdot p_i$	0.820	0.649	0.568	0.517	0.476	0.446	0.413	0.403	0.391	0.369	

Curbele de supravieţuire se ridică pe baza acestor valori p_i.

Aranjarea datelor sub forma unui tabel de acest tip este foarte convenabilă în studiile de acest gen.

b. Metoda actuarială

Faptul ca practic nu dispunem de un lot pentru a începe un studiu de tip "cohort-prospectiv" în analiza supravieţuirii impune colectarea datelor pe măsură ce apar noile cazuri (vezi fig. II.29.). Pentru efectuarea calculelor vom considera o nouă origine a timpului - în exemplul anterior a fost data primului tratament; toate intervalele se calculează în funcţie de acest moment considerat 0 pentru fiecare individ. Metoda de calcul în funcţie de această origine se numeşte **metodă actuarială.**

Pentru rata cumulativă a supravieţuirii se poate calcula şi eroarea standard conform relaţiei lui Greenwood:

$$S_p = p_n \sqrt{\sum_{i=1}^{n} \frac{q_i}{n_i p_i}} \qquad \text{(II.7.10)}$$

De exemplu, pentru *n = 10* ani de supraviețuire

$$S_p = 0{,}369\left(\frac{0{,}180}{998 \cdot 0{,}820} + \frac{0{,}209}{814 \cdot 0{,}791} + ... + \frac{0{,}057}{194 \cdot 0{,}943}\right) = 0{,}017$$

Deci cu nivel de încredere de 95%, intervalul pentru probabilitatea de a supraviețui 10 ani va fi:

$p_{10} \in (0{,}369 - 2*0{,}017 ; 0{,}369 + 2*0{,}017) = (33{,}5\% ; 40{,}3\%)$

Tabelele de viață construite după modelul tabelului II.23 sunt adaptate pentru metoda actuarială care este mai exactă decât așa numita "metodă directă" în care apar doar rapoartele privind supravițuirea pe un interval larg (5 ani, 10 ani).

c. Corectarea ratelor de supraviețuire

Concluziile pentru interpretarea ratelor de supraviețuire se obțin prin comparație, fie între diferite grupe de vârstă, fie cu rata generală de supraviețuire. În calcule pentru perioade îndelungate sau cuprinzând și pacienți mai în vârstă este recomandabil a se face corecții în raport cu rata generală de supraviețuire.

Dacă notăm cu P_0 rata generală de supraviețuire în populația generală (grupele de vârstă din care este extras lotul studiat), calculată în funcție de decesele din toate cauzele, atunci rata corectată (într-o primă aproximație) pentru supraviețuirea pe *n* ani este:

$$p_n^* = p_n / p_0 \qquad \text{(II.7.11)}$$

Valoarea lui p_0 se poate obține pentru orice țară din tabele generale de mortalitate.

d. Curbe Kaplan-Mayer

Cea mai sugestivă formă de prezentare a rezultatelor unui studiu de supraviețuire îl constituie reprezentarea grafică, în funcție de timp a ratei cumulate de supraviețuire $p_i = f(i)$ sau a ratei cumulate de mortalitate $q_i = 1 - p_i = g(i)$, cunoscute sub numele de curbe Kaplan-Mayer. În figura II.30 sunt redate aceste curbe pentru exemplul din tabelul II.23.

e. Teste

Pentru compararea a două rate de supraviețuire se pot folosi diverse teste statistice, (fie cele corespunzătoare comparării proporțiilor, fie testele t sau Wilcoxon).

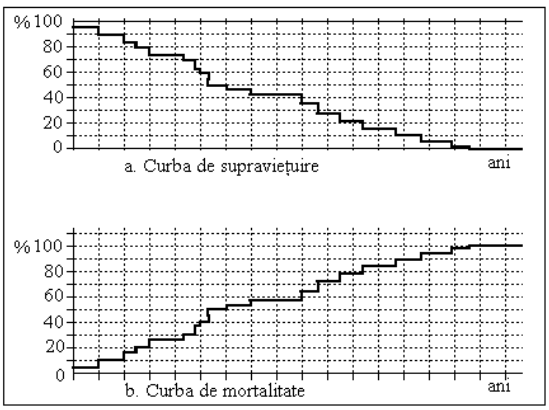


Figura II.30. Curbele Kaplan-Mayer

C. Aplicații

Analizele de tip actuarial, elaborate inițial pentru companiile de asigurări au fost extinse pentru numeroase alte situații ce implică urmărirea unei terapii: prelucrări dentare, implant cardiac, transplant de rinichi, diverse alte tipuri de protezare, etc.

În ultimul timp s-au eleborat şi modele teoretice utilizate pentru simularea fenomenelor reale, în această direcție fiind cunoscut modelul lui Cox care foloseşte o funcție "hazard" pentru descrierea matematică a ratei de mortalitate, sau modelul Cormack - Mc Kendrik pentru răspândirea epidemiilor.

BIBLIOGRAFIE SI REFERINTE

DC Altman. *Practical statistics for medical research*. Chapman&Hall/CRC, Boca Raton, 1999

P Armitage, G Berry. *Statistical methods in medical research* (2nd Ed.). Blackwell Scientific Publications, Oxford, 1987

RG Knapp, M Clinton Miller: *Clinical epidemiology and biostatistics*. Williams & Wilkins, Baltimore, 1992

DJ Sheskin. *Handbook of parametric and nonparametric statistical procedures* (3rd Ed.). Chapman & Hall/CRC, Boca Raton, 2004

Tabele de distribuții statistice: *http://www.statsoft.com/textbook/sttable.html*

Wikipedia. Teste statistice neparametrice (inclusiv tabele cu valorile critice): *http://en.wikipedia.org/wiki/Nonparametric_test*

Partea a III-a

SEMNALE ŞI IMAGINI BIO-MEDICALE

1. PRELUCRAREA SEMNALELOR BIOLOGICE

INTRODUCERE

Unul dintre capitolele cele mai bine dezvoltate ale informaticii medicale îl constituie cel referitor la prelucrarea semnalelor biologice. Funcționarea oricărui organism viu este însoțită de o permanentă modificare în timp a unor parametri (bio)fizici şi (bio)chimici. Determinarea şi înregistrarea acestor parametrii a condus la rezultate ştiințifice importante, fiind elaborate o serie de metode de investigare bazate pe culegerea lor.

Dezvoltarea deosebită a acestor metode este în bună măsură şi datorată finanțării unor astfel de cercetări de către firmele producătoare de aparatură medicală. Astăzi nici nu se mai produc electrocardiografe sau electroencefalografe fără modelul de prelucrare computerizată. În acest capitol vom trece în revistă principalele aspecte referitoare la prelucrarea semnalelor biologice, pornind de la achiziția şi filtrarea semnalelor şi analizând apoi principalele tipuri de prelucrări ale semnalelor quasi-periodice (ECG) şi neperiodice (EEG).

1.1. SEMNALE BIOLOGICE

1.1.1. Definiție. Fazele prelucrării unui biosemnal

După cum am menționat mai sus, investigația medicală modernă cuprinde urmărirea evoluției în timp a unor parametri (bio)fizici sau bio(chimici). Vom numi ***semnal biologic*** evoluția în timp a unei mărimi biologice.

Culese în forma lor naturală, semnalele biologice sunt însoțite de o serie de zgomote pe care dorim să le înlăturăm, iar din semnalul astfel curățat dorim să extragem *informația conținută de semnal reprezentată prin* parametrii relevanți pentru a caracteriza procesul generator al semnalului şi care să fie utili în decizia medicală. Putem distinge astfel principalelele faze ale prelucrării unui semnal biologic (fig. III.1):

- culegerea (achiziția) semnalului
- prelucrarea (transformarea, reducerea) semnalului
- calculul parametrilor caracteristici
- clasificarea sau interpretarea semnalului, cu scop diagnostic.

Primele două faze se referă la "sintaxa" semnalului, adică depistarea componentelor elementare ale semnalului şi urmărirea succesiunii acestora iar ultimele două se referă la "semantica" semnalului, adică semnificația acestor componente (individual sau grupate).

1.1.2. Clasificarea semnalelor biologice

Sunt posibile mai multe clasificări ale semnalelor biologice, din diferite puncte de vedere.

***a) După natura semnalului*:**

- biosemnale electrice, de ex:

= semnalul electrocardiografic ECG, datorat activității electrice a inimii

= semnalul electroencefalografic EEG, datorat manifestărilor electrice ale activității creierului

= semnalul electromiografic EMG, datorat fenomenelor electrice ce însoțesc activitatea musculară etc.

- biosemnale neelectrice de ex.:

= fonocardiograma, înregistrată din manifestările sonore ce însoțesc ciclul cardiac

= semnalul Doppler, reprezentat de variația frecvenței ultrasunetelor reflectate de suprafețe în mișcare etc.

În principiu, orice mărime biologică a cărei evoluție în timp prezintă importanţă (temperatură, pH, concentrația unor ioni etc.) poate fi considerată semnal biologic şi poate fi suspus unor metode de prelucrare cu ajutorul calculatoarelor.

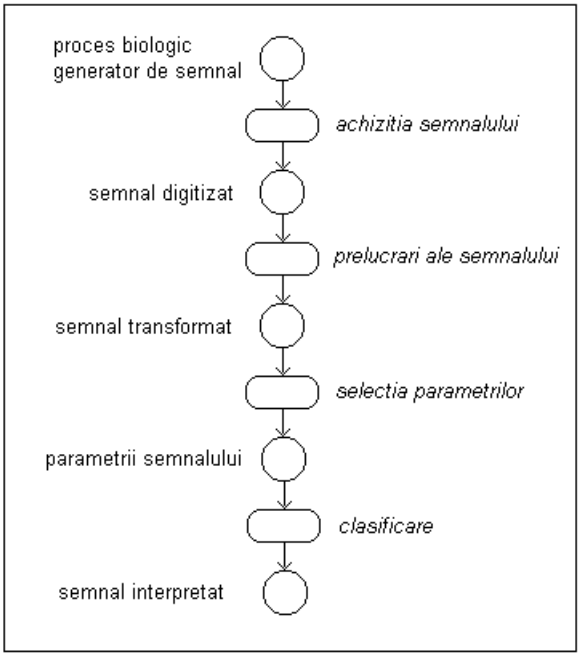


Figura. III. 1. Fazele prelucrării semnalelor biologice. Cercurile reprezintă "forma" în care este prezentat semnalul într-o anumită fază. Dreptunghiurile reprezintă fazele de prelucrare (tipuri de programe)

b) După evoluția în timp

Semnale deterministe (comportarea semnalului la orice moment poate fi predeterminată)

- semnale periodice: semnale sinusoidale armonice (fig. III 2.a)
- semnale cvasiperiodice: (de exemplu ECG) în care o succesiune de evenimente se repetă cu o anumită periodicitate (fig.III 2.b)
- semnale tranzitorii: (de exemplu potențialul de acțiune celular) care apare numai la stimulare; forma este aceeaşi ori de câte ori repetăm stimularea

Semnale stochastice (sau aleatoare): - valoarea semnalului la un moment dat nu poate fi determinată din valorile în momentele anterioare.

- semnale staționare: (de exemplu EEG) în care anumiți parametri (de exemplu media) rămân constanți (fig.III. 2.d)

- semnale nestaţionare: (de exemplu EMG) în care şi parametrii statistici depind de timp (fig. III. 2.e).

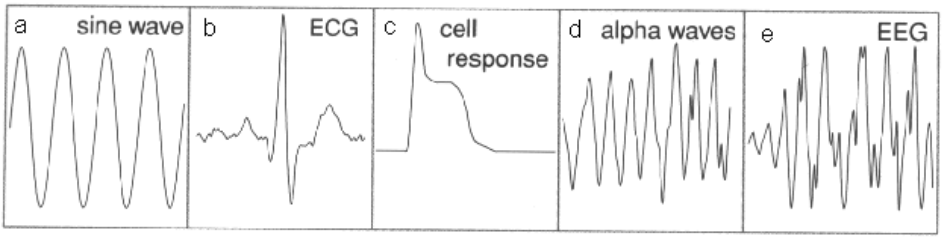


Figura III.2. Clasificarea semnalelor biologice: deterministe: a, b, c, şi stochastice: d, e. Subdiviziuni: **a**: semnal periodic (undă sinusoidală); **b**: semnal cvasiperiodic (ECG); **c**: semnal transzitoriu (potenţial de acţiune); **d**: semnal aleator staţionar (fusuri alfa în EEG); **e**: semnal aleator nestaţionar (EEG). [Van Bemmel&Musen 1997]

O categorie aparte de semnale o constituie aşa-numitele *"trenuri de impulsuri"* (procese punctiforme - "point processes") în care nu ne interesează forma semnalului ci numai apariţia sau nu a unui impuls (exemplu: apariţia undelor R în semnalul ECG sau impulsurile nervoase pe axoni); aceste semnale se descriu prin aşa-numitele impulsuri Dirac - funcţii care au valoarea zero peste tot, exceptând intervalele foarte scurte în care apar evenimentele.

1.1.3. Electrozi de culegere. Traductori

Semnalele de natură electrică (ECG, EEG, EMG, etc.) reprezentând manifestări electrice ale fenomenelor studiate (ale inimii în EEG, ale creierului în EEG, ale muşchiului în EMG) sunt culese cu ajutorul unor *electrozi* puşi în contact cu ţesutul analizat sau – cel mai adesea – pe piele în regiuni în care se proiectează aceste activităţi electrice. De obicei aceşti electrozi sunt confecţionaţi dintr-un metal impolarizabil (Ag), acoperiţi cu un tifon umezit cu soluţie salină sau gel conductor, pentru a asigura un bun contact electric. Cel mai adesea aceşti electrozi de culegere sunt menţinuţi în poziţia de contact cu ajutorul unor benzi de cauciuc. Este bine a se acorda atenţie fixării acestor electrozi şi asigurarea unui contact bun pentru a evita o întreagă gamă de artefacte posibile. Asistentele experimentate depistează rapid electrozii plasaţi incorect. Semnalul electric cules de aceşti electrozi este filtrat şi amplificat fiind în continuare supus operaţiilor de prelucrare.

În cazul semnalelor care nu sunt de natură electrică ci de altă natură (mecanică: contracţii, chimică: concentraţii etc.), se folosesc dispozitive numite ***traductori*** care transformă semnalul original în semnal electric.

Actualmente s-au realizat traductori care pot transforma în semnal electric aproape orice tip de mărime: presiune, forţă, temperatură, deplasare, pH, concentraţia unei substanţe (în ultimul timp s-au realizat "biosenzori" pentru unele molecule organice) etc. Există în momentul de faţă o adevărată cursă pentru realizarea unei palete largi de biosenzori, pe de o parte pentru comoditatea de lucru comparativ cu metodele chimice (răspuns rapid, fără manevre suplimentare, suficient de precis), pe de altă parte pentru posibilitatea urmăririi în timp a parametrilor pentru o perioadă mai îndelungată de timp.

Semnalele bioelectrice au în general valori foarte mici (milivolţi, chiar microvolţi) şi de aceea trebuiesc **amplificate** înainte de începerea prelucrării.

1.2. ACHIZIȚIA BIOSEMNALELOR

1.2.1. Sisteme de culegere a biosemnalelor

După cum am menționat mai sus, semnalele biologice pot fi culese în două moduri:

a) în cazul în care biosemnalul nu este de natură electrică se folosec **traductori** care transformă semnalul original în semnal electric.

b) în cazul în care biosemnalul este electric se folosec pentru culegere nişte **electrozi de culegere** care pot fi:

- *electrozi superficiali* (de exemplu în ECG, EEG), care se aplică la suprafața tegumentelor; pentru a asigura o bună conducere electrică suprafețele unde se aplică electrozii (de obicei argintați, acoperiți cu tifon) se umezesc cu soluție salină sau gel salin.
- *electrozi-ac*, de exemplu în EMG, în cadrul metodelor invazive.

1.2.2. Conversia analog-numerică

Semnalul cules şi amplificat se prezintă uzual ca *o succesiune continuă în timp* a unor diferențe de potențial, fiind deci un *semnal analogic*. Pentru a putea prelucra un semnal cu ajutorul unui calculator numeric este necesar a transforma semnalul analogic (continuu), într-o succesiune de valori numerice, care reprezintă un *semnal numeric* (sau *digital*). Transformarea se face prin "citirea" valorilor semnalului real (continuu) la anumite intervale de timp.

Definiție. *Transformarea unui semnal analogic (continuu) în semnal numeric discret (digital) se numeşte* ***conversie analog-numerică (digitală)***.

Deci descrierea numerică a unui semnal este o descriere discontinuă (discretă). Dacă intervalul de timp între două "citiri" succesive este suficient de scurt, descrierea este fidelă. Realizarea conversiei analog-numerice cuprinde două elemente fundamentale: **eşantionarea şi cuantizarea**. Încadrarea convertorului analog-numeric (CAN) în schema de achiziție a unui semnal este prezentă în fig. III.3.

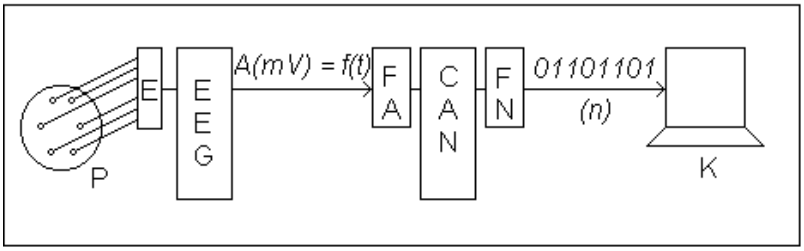


Figura III.3. Plasarea convertorului analog numeric CAN în sistemul de achiziție a semnalului EEG între pacientul P şi calculatorul K. Sistemul de conexiune a electrozilor E este legat de electroencefalograf EEG. Semnalul analogic exprimat de amplitudinea A(μV) funcție de timp este "digitizat", devenind semnal numeric exprimat pe **n** biți. Dacă s-ar folosi filtre, atunci filtrul analogic FA s-ar plasa înainte de CAN iar cel numeric FN după CAN

a) Eşantionarea semnalelor

Operația de discretizare a axei orizontale (abscisa) a unui semnal se numeşte ***eşantionare.*** Cum pe abscisă noi reprezentăm timpul, putem spune că eşantionarea reprezintă "citirea" semnalului la intervale discrete de timp. Putem astfel defini două mărimi caracteristice ale eşantionării: perioada de eşantionare şi frecvența de eşantionare.

Perioada de eşantionare *a unui semnal reprezintă intervalul de timp între două citiri succesive ale valorilor semnalului. Numărul de citiri ale semnalului în unitatea de timp se numeşte* ***frecvență de eşantionare.***

Dacă notăm perioada de eşantionare cu T_e şi frecvența de eşantionare cu f_e, atunci:

$f_e = 1 / T_e$ sau $T_e = 1/f_e$ (III.1)

Când perioada de eşantionare se exprimă în secunde (s) obținem frecvența de eşantionare în herzi (Hz).

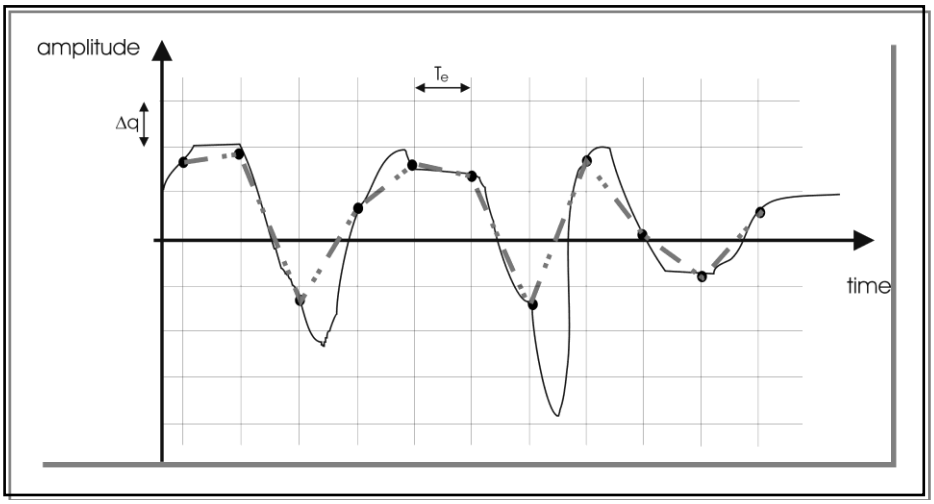


Figura III.4. Conversia analog-numerica: eşantionarea şi cuantizarea, cu linie continuă este reprezentat semnalul real; cu linie intreruptă este reprezentat semnalul eşantionat. Cunoaştem doar valorile măsurate în punctele de citire (la intervalele de timp date de perioada de eşantionare Te). Distanța Δq dintre două trepte de amplitudine determină precizia citirii valorilor

Un exemplu de eşantionare greşită a semnalului este prezentat în fig. III.5. stânga. Se observă că perioada de eşantionare aleasă este prea mare; avem o variație atât de rapidă a semnalului încât ea va trece neobservată. De aceea, pentru a urmări semnalul real cu fidelitate trebuie să alegem o perioadă de eşantionare foarte scurtă, deci o frecvență de eşantionare ridicată. Dacă notăm frecvența maximă a semnalului (numită şi frecvență Nyquist) cu f_{max} atunci frecvența de eşantionare trebuie să respecte condiția (III.2):

$$f_e \geq 2*f_{max} \quad \text{(III.2)}$$

Această condiție se mai numeşte "teorema de eşantionare" sau teorema Shannon-Nyquist, care se enunță astfel:

Frecvența de eşantionare trebuie să fie cel puțin dublă față de frecvența maximă a semnalului.

În figura III.5.dreapta s-a crescut frecvența de eşantionare şi nu se mai pierd detaliile privind variațiile semnalului.

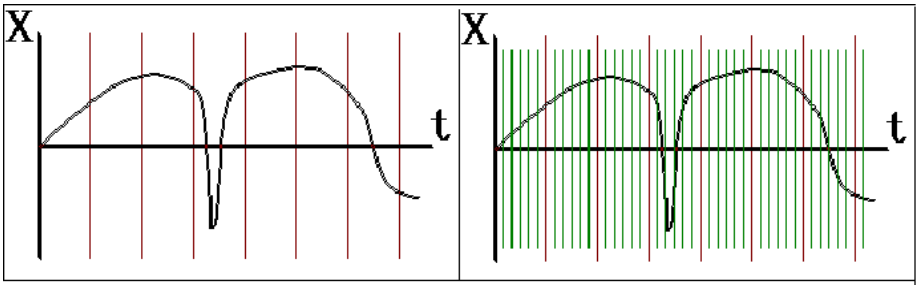


Figura III.5. Ilustrarea teoremei de eşantionare

Am fi tentaţi să credem că este foarte bine să luăm o frecvenţă de eşantionare cât mai mare, însă această creştere duce la achiziţionarea unui număr ridicat de valori numerice pentru acelaşi interval de timp, determinând o creştere substanţială a timpului de prelucrare (fără a obţine totdeauna o creştere semnificativă a calităţii rezultatelor). De aceea frecvenţa de eşantionare se alege la dublul frecvenţei Nquist sau uşor peste această valoare.

Uzual se folosesc frecvenţe de eşantionare de 60-100 Hz (EEG), 250-500 Hz (ECG), până la 1-10 kHz (EMG; potenţiale evocate).

b) Cuantizarea semnalelor

Intervalul de valori cuprins între valorile extreme posibile (minimă şi maximă) ale semnalului se împarte într-un număr **N** de **trepte de amplitudine**, astfel încât practic se citesc valorile corespunzătoare treptelor date de Δq (fig III.4). Cu cât numărul de trepte este mai mare, cu atât precizia de citire este mai bună.

Valorile citite se exprimă în sistem binar. De aceea, cel mai adesea numărul treptelor de amplitudine este o putere a lui 2. De exemplu, pentru 256 trepte, o valoare citită este exprimată pe 8 biţi, căci $256 = 2^8$. Se obişnuieşte să se caracterizeze un convertor analog-numeric prin **numărul de biţi** prin care se reprezintă o valoare citită. În majoritatea tipurilor de prelucrări întâlnite la analiza semnalelor biologice se foloseşte o reprezentare pe 12 biţi, (uneori sunt suficienti 8 sau 10 biţi). Mai rar (potenţiale evocate, EMG) se folosesc convertoare pe 16 biţi.

Relaţia între **N** - numărul treptelor de amplitudine (cuantizare) şi **n**- numărul de biţi prin care se exprimă valoarea citită de CAN este:

$$N = 2^n \qquad \text{(III.3)}$$

Putem astfel exprima sensibilitatea de citire a CAN, adică variaţia potenţialului de intrare care corespunde unei modificări de 1 bit a valorii citite. Această sensibilitate se mai numeşte rezoluţie de amplitudine sau precizie de citire sau **cuantă de citire** şi are valoarea:

$$\Delta V = \frac{V\max - V\min}{N} = \frac{V\max - V\min}{2^n} = \Delta q \qquad \text{(III.4)}$$

unde:

V_{max}, V_{min} sunt valorile extreme posibile ale semnalului
N - numărul de trepte de amplitudine (cuantizare)
n - numărul de biţi ai CAN.

1.2.3. Multiplexarea

De obicei înregistrarea biosemnalelor se realizează folosind mai mulți electrozi de culegere care se aranjează în diferite moduri numite **derivații**, cel mai adesea standardizate. Fiecare electrod culege semnalul pentru un **canal**. Echipamentele de conversie analog-numerică permit înregistrarea pe mai multe canale folosind un singur convertor care este comutat pe rând la toate canalele cu ajutorul unui dispozitiv numit "**multiplexor**". În cadrul programelor trebuie să se țină cont de întârzierea dintre citirile efectuate pe diferite canale. Există şi multiplexoare care citesc valorile pe toate canalele aproape în acelaşi moment (cu o frecvență de eşantionare foarte ridicată); după o pauză urmează o nouă "salvă" de citiri.

1.3. SPECTRE DE FRECVENȚĂ ŞI FILTRARE

1.3.1. Reprezentarea semnalelor

Semnalele se pot reprezenta ca o funcție de timp [ampl.= f(timp) - evoluția în timp a unei mărimi] sau ca o funcție de frecvență [ampl. = f(frecvența) sau putere = f(frecvența) - punând în evidență compoziția semnalului]. Reprezentările în funcție de frecvență se numesc *spectre de frecvență*. Figura III.6 arată spectrele de frecvență ale unor semnale periodice uşor de identificat (A, B şi C) şi a unui semnal neperiodic (D).

1.3.2. FILTRAREA BIOSEMNALELOR

Zgomote

Semnalele bioelectrice au în general valori foarte mici iar acțiunea de culegere a lor este însoțită de culegerea unor zgomote care perturbă (uneori foarte puternic) semnalul. Pentru a îmbunătăți raportul între semnalul util şi zgomot, odată cu amplificarea semnalului se realizează şi o **filtrare** pentru eliminarea zgomotelor. Pentru a putea înlătura în mod specific zgomotele (parțial sau total) să urmărim o clasificare a lor.

Clasificarea zgomotelor se poate face din mai multe puncte de vedere.

a) După evoluție:

- zgomote **(cavsi)periodice**, numite şi zgomote "roz", în care sunt predominante anumite frecvențe,
- zgomote **neperiodice**, numite şi zgomote "albe" în care frecvențele componente au aceeaşi probabilitate.

b) După tendință:

- zgomote **sistematice** - de exemplu cele datorate unui electrod plasat necorespunzător,
- zgomote **întâmplătoare.**

c) După origine:

- la culegere - datorate amplasării nepotrivite a electrozilor sau unor contacte electrice nesatisfăcătoare
- la amplificare - majoritatea amplificatoarelor amplifică neuniform diferite domenii de frecvență; ponderea acestor zgomote este destul de redusă, aparatele moderne având amplificatoare de bună calitate,
- artefacte "bio" - se întâmplă uneori ca, pe lângă semnalul dorit a se înregistra, să fie culese şi alte semnale (de exemplu semnal electrocardiografic suprapus peste EEG sau artefactele de respirație în înregistrarea ECG).

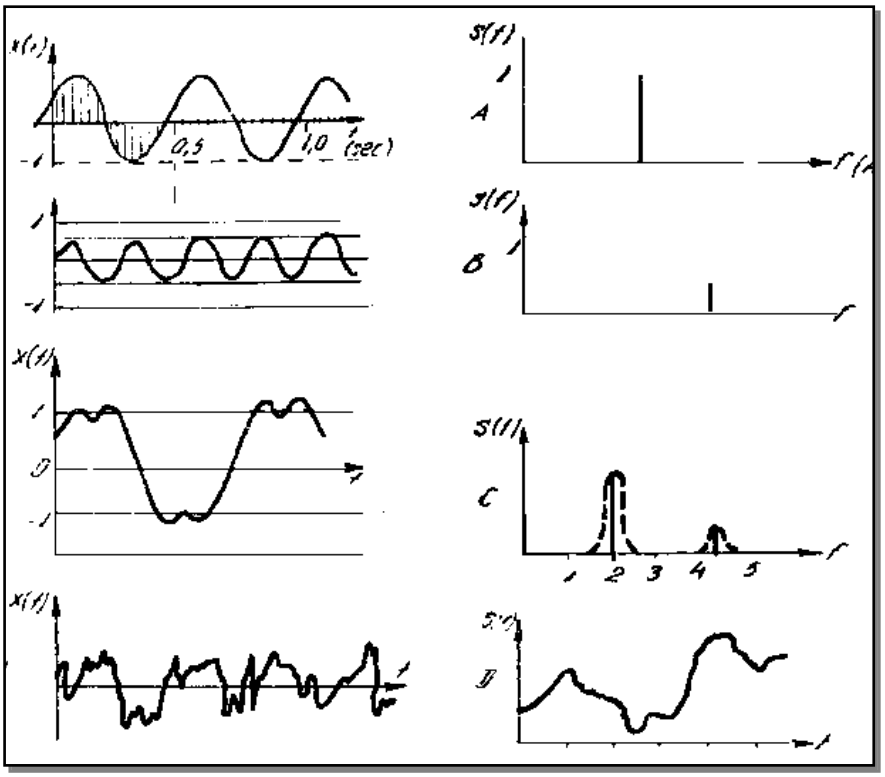


Figura III.6. Spectrele diferitelor tipuri de semnale: a) Un semnal sinusoidal cu frecvența *f = 2Hz* și spectrul său în A. b) Un semnal de *4,5 Hz* cu amplitudine mai mică și spectrul său în B. c) Semnalul rezultat din suprapunerea semnalelor din a) și b); spectrul său în C cuprinde două linii (cu linie întreruptă este prezentat spectrul când se prelucrează un tronson mai scurt din semnal). d) Pentru un semnal neperiodic, spectrul (D) este continuu. [Popescu 1988]

Tipuri de filtre

a) După regiunea admisă (figura III.7)

- filtru "**trece sus**" (fig.III.7.a) care lasă să treacă toate frecvențele $f \geq f_o$
- filtru "**trece jos**" (fig. III.7.b) care lasă să treacă numai frecvențele $f \leq f_o$
- filtru "**trece bandă**" (fig. III.7.c) care lasă să treacă frecvențele cuprinse între două limite: $f_i \leq f \leq f_s$
- filtru "**opreşte bandă**" (fig. III.7.d) are frecvențele $f \leq f_i$ și $f \geq f_s$
- filtru "**ac**" de tip "opreşte" (sau "trece") în care regiunea dintre cele două limite f_i - f_s este foarte îngustă; se utilizează în special pentru eliminarea perturbațiilor produse de posturile locale de radio;

b) filtre analogice - filtre numerice: cele analogice sunt utilizate ca dispozitive fizice înainte de intrarea semnalului în convertorul analog-digital, în timp ce filtrele numerice se aplică semnalului deja digitizat (fig. III.3 arată plasarea lor);

c) filtre fără memorie - cu memorie: cele fără memorie au ca secvență de ieşire o sumă ponderată a unei perioade finite de intrare si au avantajul unei ieşiri identice pentru aceeași intrare; filtrele cu memorie țin cont de un număr de ieşiri anterioare și au avantajul de a folosi un număr redus de coeficienți pentru ieşire, însă necesită o inițializare;

d) filtre nerecursive - recursive: în cazul filtrelor digitale calculul coeficienților pe cale recursivă este redus;

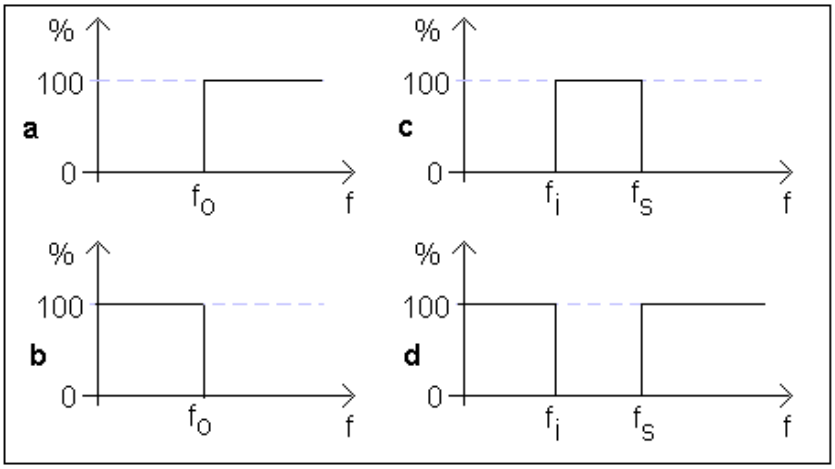


Figura III.7. Tipuri de filtre – clasificare dupa regiunea admisă: anumite frecvenţe sunt lăsate să treacă, pe când altele sunt înlăturate

e) filtre generale - dedicate: cunoscând caracteristicile semnalului, raportul semnal-zgomot este îmbunătăţit în filtrele construite special pentru domeniul de frecvenţe şi amplitudini dorit, precum şi pentru tipul de unde; deoarece în multe semnale biologice apar atât fenomene mai lente, chiar filtrele special construite pentru un anumit tip de semnal păstrează un caracter mai general;

f) filtre invariante în timp - filtre adaptive: construcţia unui filtru ale cărui caracteristici frecvenţiale să se adapteze semnalului necesită un semnal de referinţă, care în unele situaţii poate fi generat - fiind astfel posibil să se suprime unele interferenţe nedorite în semnal;

g) filtre liniare - neliniare: se definesc în funcţie de relaţia diferitelor componente în structura spectrală a semnalului de ieşire, funcţie de cea de intrare.

1.4. PRELUCRAREA SEMNALELOR CVASI – PERIODICE. SEMNALUL ELECTROCARDIOGRAFIC

Semnalele cvasi-periodice, dintre care semnalul electrocardiografic (ECG) este cel mai reprezentativ, necesită o prelucrare în care se porneşte de la detecţia perioadei, urmată de detecţia unor evenimente în cadrul perioadei şi caracterizarea parametrică a undelor şi/sau intervalelor.

1.4.1. Semnalul ECG

Semnalul ECG reprezintă un semnal electric de mică amplitudine ce reflectă la nivel superficial activitatea electrică a inimii. Inima este un organ ce reprezintă un automatism funcţional. Declanşarea unei revoluţii cardiace începe printr-o depolarizare a nodului sino-atrial care se propagă la nodul atrioventricular. Această depolarizare este reprezentată în traseul ECG (vezi figura III.8) prin unda P. Unda de depolarizare se propagă generând depolarizarea ventriculară reprezentată de complexul QRS, urmat de repolarizare reprezentată de unda T. În cazuri patologice se observă diverse modificări, prelucrarea cu calculatorul având scopul de a creşte sensibilitatea sesizării acestor modificări şi a realiza clasificarea lor.

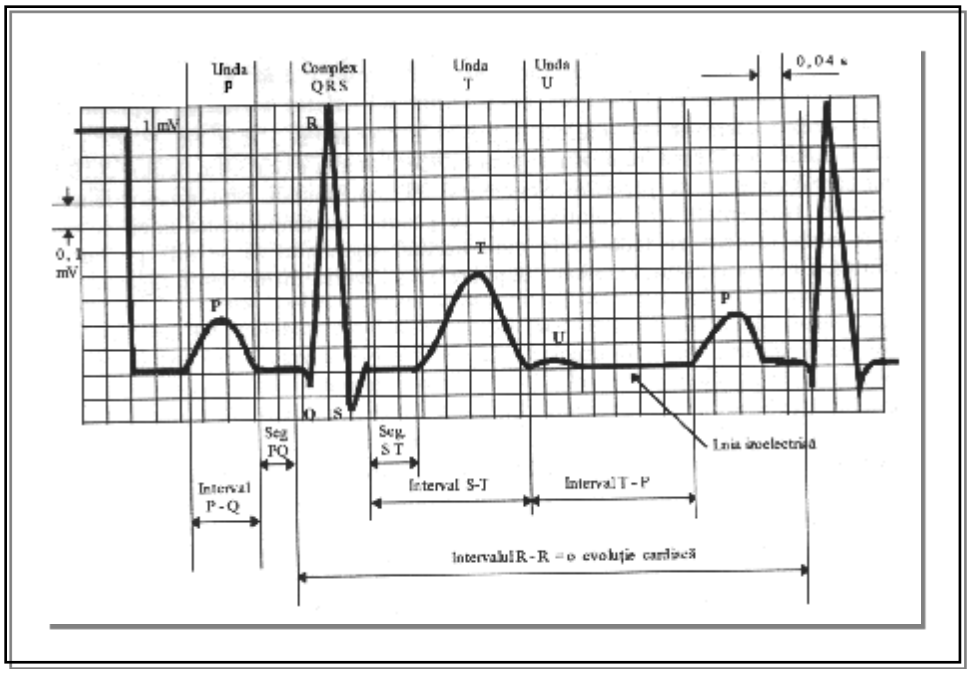


Figura III.8. Traseul ECG normal

1.4.2. Achiziția semnalului ECG

Ca pentru orice semnal, conversia analog-digitală, conformă teoremei de eşantionare a lui Shannon, preia semnalul filtrat (în general filtre "trece" de bandă 0,5-40 Hz), apoi îl eşantionează (actualmente se foloseşte f_e=250 - 500 Hz) şi îl cuantizează pe 8, 10 sau 12 biți.

Rezultatele obținute cu o frecvență de eşantionare constantă sunt destul de bune, însă, datorită faptului că ritmul cardiac nu este constant, se poate întrebuința şi o frecvență de eşantionare adaptabilă la ritm, astfel ca fiecare bătaie a inimii să fie împărțită în acelaşi număr de puncte.

Dezvoltarea unor aplicații care să sesizeze toate abaterile posibile ale semnalului de la normal a concentrat însemnate eforturi, ilustrate într-o bogată literatură consacrată acestei teme:

- programe pentru interpretarea ECG în cele 12 derivații
- programe pentru vectocardiograme
- programe pentru ECG şi VCG
- pograme pentru ECG şi VCG în efort.

1.4.3. Detecția perioadei

În cadrul semnalelor (cvasi-)periodice este esențială detecția perioadei - intervalul de timp după care se repetă acelaşi ciclu de evenimente.

Din punct de vedere funcțional perioada este definită ca intervalul între începutul cascadei de evenimente care debutează cu depolarizarea nodului sino-atrial (începutul undei P) şi începutul următorului ciclu, adică intervalul între două unde P (fig. III.9). Unda P, având amplitudine mică (0,1 - 0,2 mV), este detectată destul de greu de catre algoritmii utilizati în aplicațiile de prelucrare automată. De aceea aceste programe utilizează ca repere unde de amplitudini mari, de exemplu unda R (cca 1mV). Se defineşte astfel perioada detectabilă R-R.

Metoda uzuală de detecţie se numeşte "metoda intersecţiei de nivel". Se alege un nivel de referinţă (de ex. 0.9 mV) şi se compară fiecare punct cu această valoare; se reţin indicele punctului la o primă traversare a nivelului şi cel al traversării următoare. Cunoscând frecvenţa de eşantionare şi cei doi indici se calculează imediat intervalul de timp între două bătăi.

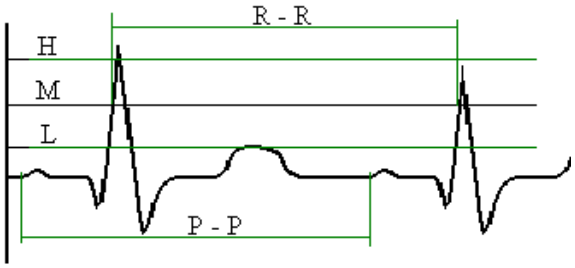


Figura III.9. Detecţia perioadei prin metoda intersecţiei de nivel

Nivelul de intersecţie trebuie ales corespunzător. Pentru a fi cât mai aproape de vârful undei R s-ar recomanda o valoare ridicată, însă există riscul de a "pierde" bătăi, deoarece nu toate vârfurile R au aceeaşi amplitudine (nivelul H în fig. III.9). În plus, să nu uităm că este vorba de un semnal eşantionat şi rareori citim chiar valoarea de vârf. Pe de altă parte, la coborârea nivelului de intersecţie se reduce (până la anulare) riscul de a "pierde" bătăi cardiace, dar apare riscul intersecţiei cu unda T care ajunge uneori la aproape jumătate din amplitudinea undei R şi va fi deci interpretată ca o nouă bătaie (nivelul L în fig. III.9). De obicei se estimează amplitudinea a cca 10 vârfuri R succesive, şi se ia ca nivel de intersecţie o valoare de 75-80% din media acestor vârfuri (nivelul M în fig. III.9).

În programele actuale detecţia perioadei se realizează foarte exact. Există numeroase programe care reţin aceste perioade pe o durată îndelungată făcând posibilă reprezentarea grafică a evoluţiei pulsului pacientului.

1.4.4. Etapele interpretării ECG

Deşi programele enumerate diferă prin unele caracteristici, toate urmăresc anumite elemente fundamentale, prin *reducerea şi transformarea semnalului într-un set de câţiva parametri semnificativi* pentru deciziile ulterioare. Etapele, indiferent de metodele specifice prin care se realizează cuprind:

a) detecţia complexelor QRS
b) detecţia artefactelor:
- corecţia liniei de zero
- artefacte musculare

c) tipificarea complexelor QRS
d) tipificarea ST-T
e) detecţia undelor P
f) selecţia şi medierea ciclurilor
g) recunoaşterea undelor
h) recunoaşterea *pattern*-ului, cuprinzând:
- extragerea parametrilor
- clasificarea - programe de diagnostic.

Majoritatea sistemelor utilizate divid prelucrarea în etapele menţionate, fiecare etapă reprezentând un modul program ce conţine un set independent de subrutine.

Modulele sunt conectate la dispozitivul de stocare a datelor printr-o interfață *software*. Sistemele sunt, în general, independente de frecvența de eşantionare, care variază între 250 şi 500 Hz. De asemenea, modulele trebuie să fie independente de derivațiile analizate. În majoritatea cazurilor, cele 12 derivații sunt împărțite în 4 grupe de câte 3 derivații culese simultan.

1.4.5. Descrierea modulelor de prelucrare ECG

Din punct de vedere al informaticianului, sarcinile ce trebuie rezolvate în prelucrarea ECG determină împărțirea în module:

1. *introducerea datelor:* este un modul ce dirijează convertorul la culegerea datelor, fie pe o bandă magnetică, fie *on-line* de la pacient
2. *detecția QRS:* este primul pas în toate sistemele; fiecare derivație este inspectată pentru prezența *spike* - urilor şi se determină un punct de referință ("punct fiducial") în complexele QRS; este dificilă separarea unui QRS de un artefact puternic în vecinătatea sa
3. *detecția artefactelor:* în cazul depăşirii unor nivele, în unele derivații, datele pot fi filtrate sau eliminate în cazul nivelelor de saturație
4. *clasificarea QRS:* complexele QRS detectate sunt grupate în familii, după forma undelor, stabilindu-se tipul dominant; se măsoară şi fluctuația intervalului între complexe QRS succesive, utilă în analiza ritmului; tot acum se evaluează dacă abaterile liniei de bază depăşesc un anumit nivel
5. *tipificarea ST:* se detectează începutul segmentului ST (punctul "J" jonction), se compară între ele segmentele ST - T ale complexelor QRS dominante şi se rețin pentru mediere cele asemănătoare
6. *detecția undelor P:* este cercetată activitatea atrială, detectându-se undele P, atât cele la distanță fixă de QRS, cât şi cele la distanță variabilă; se detectează şi flutterul atrial, dacă este present
7. *modulul "bătaie":* când s-au găsit suficiente complexe QRS cu segmente ST-T asemănătoare, se mediază cu punctul de referință găsit în detecția QRS
8. *durata QRS:* se determină începutul şi sfârşitul complexelor QRS mediate; aceasta se efectuează pentru cele trei derivații simultan în fiecare grup de derivații
9. *durata P:* se poate stabili numai dacă distanța față de începutul complexului QRS este fixă; se efectuează simultan în cele trei derivații din fiecare grupare
10. *sfârşitul undei T:* stabilirea sfârşitului undelor T
11. *parametrizarea:* pentru fiecare derivație se rețin amplitudinea şi durata Q, R, S şi amplitudinile P, T (uneori şi alți parametri)
12. *ritmul:* se realizează o clasificare a ritmului, conform datelor furnizate de modulele anterioare
13. *clasificare a conturului:* se efectuează atât pentru ECG cât şi în VCG; se utilizează adesea codul Minnesota şi programul de diagnostic stabilite de sistemele IBM şi Mayo Clinic
14. *prezentarea rezultatelor:* modulul de ieşire cuprinzând unii parametri, graficul complexelor mediate şi diagnosticul.

După cum se observă, modulele 2, 3, 6, 8, 9, 10 se referă la probleme tipice de detecție, restul fiind module de recunoaştere a formelor (11 de tip extragerea atributelor, iar 4, 5, 12, 13 de clasificare).

1.4.6. Descrierea etapelor de detecție

a) Detecția QRS poate fi efectuată prin algoritmi pentru câte o derivație sau pentru mai multe derivații. Cea mai uzuală tehnică detectează un prag stabilit. Semnalul original este filtrat (filtru trece bandă), iar fiecare traversare a pragului într-un anumit sens

este reținută; se introduce apoi selecția, prin eliminarea traversărilor la intervale prea scurte față de precedentele.

O altă metodă, propusă de Udapa şi Murthy, utilizează descrierea sintactică a complexelor ventriculare şi supraventriculare pentru analiza ritmului. Se introduc 7 simboluri pentru aprecierea fiecărui segment de eşantionare, conform pantelor (0, uşor, mediu sau puternic, pozitive sau negative); semnalul este transformat într-o propoziție. Se definesc gramatici pentru complexele ventriculare, care sunt recunoscute prin analiză sintactică (analiza propozițiilor).

În cazul analizei simultane pe mai multe derivații, în special în VCG, se definesc vectori tridimensionali (în cazul a 3 derivații) şi se introduce termenul de viteză spațială pentru variația vectorului; complexul QRS este detectat dacă viteza spațială depăşeşte un prag; deseori analiza pe mai multe derivații utilizează intervalul de suprimare folosit în analiza pe o singură derivație.

Detecția QRS fiind o problemă fundamentală în analiza semnalului ECG, programele sunt în continuă îmbunătățire, ajungându-se în momentul de față ca numărul erorilor (QRS fals pozitive şi fals negative) să fie destul de redus.

b) Detecția artefactelor. Calitatea semnalului de intrare este o condiție esențială pentru interpretarea traseului, indiferent dacă este efectuată de om sau computer. Artefactele traseelor înregistrate pot fi împărțite în 5 categorii, fiecare cu caracteristicile sale:

- devierea liniei de bază
- interferența frecvenței rețelei de curent
- artefacte musculare
- *spike*-uri
- saturația de amplitudine prin modificări bruşte ale liniei de bază.

Fiecare tip de artefact necesită metode specifice de detecție şi corecție.

. *devierea liniei de zero* este o perturbație de joasă frecvență, datorată respirației sau mişcării pacientului. Aceste devieri sunt imprevizibile şi deci greu de corectat. Sunt posibile mai multe căi de corecție, cea mai uzuală fiind **metoda lui Riedl**. Din semnalul eşantionat se selectează fiecare al 20-lea punct (care poate aparține fie liniei de bază, fie unei unde); se elimină punctele cu variații mari față de cele vecine şi se netezeşte curba, obținându-se profilul liniei de bază, ce va fi considerată linie de zero; prin scăderea ei din semnalul original se obține semnalul corectat (fig.III.10);

. *interferența frecvenței rețelei* este o perturbație previzibilă şi corecția se poate efectua prin filtrare. Este, totuşi, important de remarcat că unele componente din complexul QRS se situează în acelaşi domeniu de frecvențe cu frecvența retelei de alimentare şi astfel se produce o perturbare nedorită a semnalului original. Mortara a descris o tehnică neliniară de estimare a interferenței rețelei, bazată pe predicția semnalului la momentul ulterior; este astfel posibilă o filtrare numerică, dar complexele QRS pot fi afectate şi în această situație;

. *artefactele musculare* au un spectru mai larg de frecvență şi apar, de regula, în înregistrările în timpul efortului; o filtrare de joasă frecvență reduce, în general, contribuția lor la distorsionarea semnalului;

. *spike-urile,* prin durata lor, deosebit de scurtă, pot fi recunoscute mai uşor şi eliminate printr-o procedură de comparație cu punctele vecine; la frecvențe de eşantionare mai mici, există pericolul de confuzie cu componentele QRS;

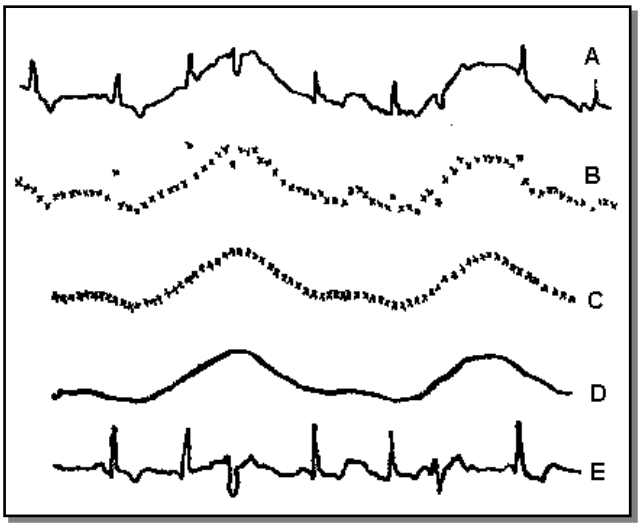


Figura III.10. Exemplu de corecție a liniei zero. [Popescu 1988]

. *saturarea în amplitudine* este tot o deviere a liniei de bază, însă bruscă; corecția se face similar, introducând condiția de eliminare prin comparația mai multor puncte vecine; în general, prin această corecție se pot pierde unele complexe QRS, deci zonele corectate se etichetează.

c) Tipificarea QRS: deşi este în mare măsură o problemă de recunoaştere a *pattern*-ului, (fig. III.11) are unele aspecte specifice:

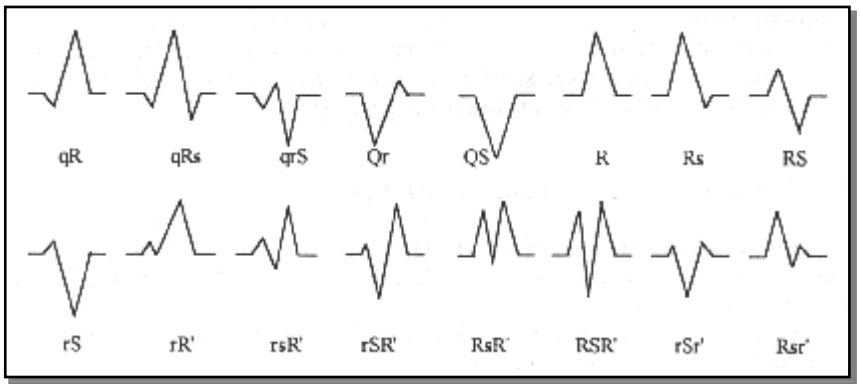


Figura III.11. Tipuri de complexe QRS. [Popescu 1988]

- alinierea complexelor QRS înainte de extragerea atributelor este necesară dacă la detecție s-a utilizat direct amplitudinea semnalului şi nu derivata sa; alinierea se realizează prin modificarea poziției relative a întregului complex în raport cu punctul de referință (metoda utilizată este maximizarea coeficientului de corelație între complexe);

- alegerea parametrilor pentru extragerea atributelor oferă o paletă largă de posibilități privind punctele de calcul ale amplitudinilor şi duratelor. Diferitele programe enumerate la începutul paragrafului utilizează seturi variate de atribute, folosind fie valori culese, fie valori de interpolare; se susține adesea că o frecvență de eşantionare ridicată (500 Hz) ar deveni preferabilă prin calitatea oferită în această prelucrare.

d) Tipificarea ST - T este, de asemenea, o problemă abordabilă în stilul clasic al recunoaşterii *pattern* - ului (fig. III.12).

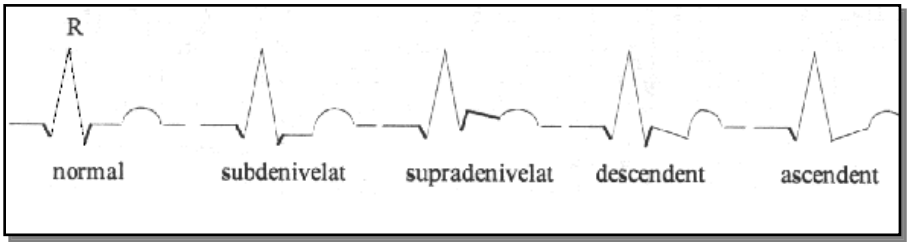


Figura III.12. Variante de segmente ST. [Popescu 1988]

Situaţii mai deosebite apar când corecţia liniei de bază a fost insuficientă; ele pot fi lesne confundate de computer cu modificări ale conducţiei ventriculare. Totuşi, dacă pentru mediere au rămas suficiente segmente ST -T, zgomotul introdus este în mare măsură suprimat.

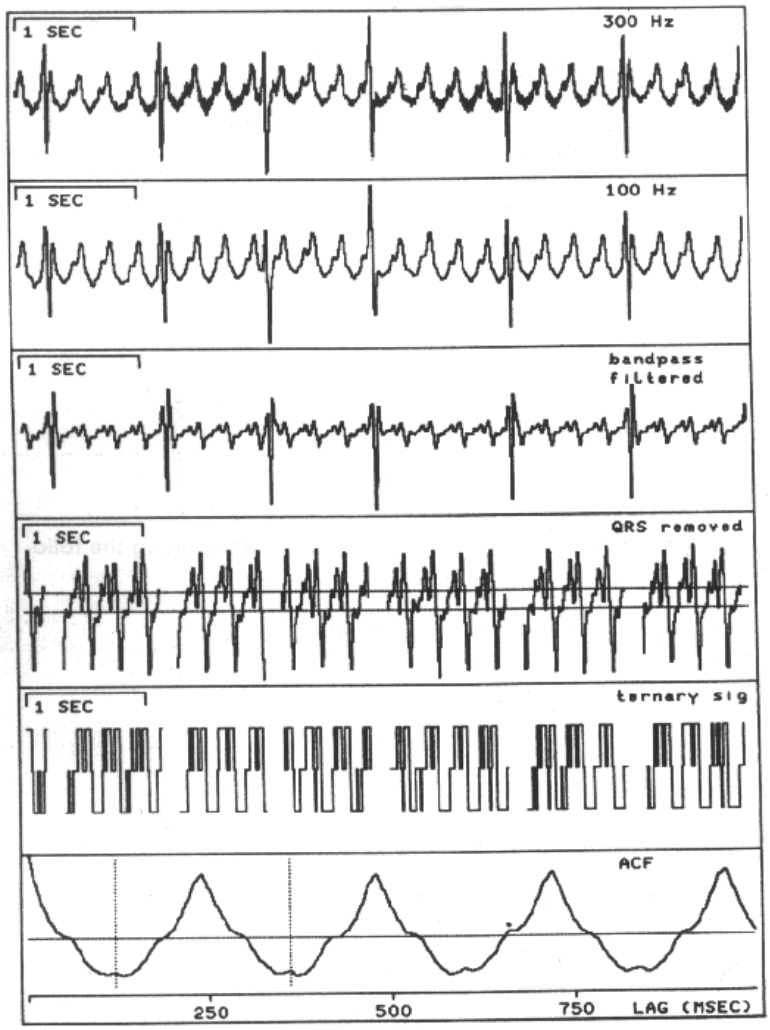


Figura III.13. Exemplu cuprinzând paşii de prelucrare a funcţiei de autocorelaţie pentru detecţia undelor *flutter*. [Popescu 1988]

e) Detecția undelor P este una dintre cele mai dificile probleme în analiza ECG sau VCG, în literatură existând încă unele neconcordanțe de terminologie, detecția undelor P fiind uneori considerată ca determinare a începutului şi sfârşitului undei P, care fiind o undă de mică amplitudine poate fi uşor confundată cu un zgomot. Este mai rezonabil a căuta întâi prezența undei şi apoi a-i delimita marginile.

Sunt detectabile 3 tipuri de activități atriale:

- unde P urmate de complexe QRS la intervale fixe, numite unde P cuplate; în aceste cazuri timpul de conducere atrio-ventriculară este aproximativ constant (variabilitate de maximum 30 ms) - ritmul sinusal
- unde P ce nu sunt întotdeauna urmate de complexe QRS, sau distanța este variabilă, ca în cazurile de bloc atrio-ventricular de gradul 2
- unde *flutter*, definite ca oscilații regulate bifazice, uniforme, cu frecvențe între 200-400 pe minut; nu se evidențiează o linie de bază.

Dacă nu apare nici unul din tipurile descrise, abaterea de la normal va fi decisă de modulul de clasificare a ritmului; astfel de situații apar în fibrilația atrială sau ritmurile nodale.

Un exemplu privind paşii de prelucrare pentru detecția undelor *flutter* prin funcția de autocorelație este redat în fig. III.13. Metoda este descrisă în paragraful privind analiza temporală a semnalului EEG.

f) Selecția de mediere a ciclurilor. Aproape toate sistemele de prelucrare utilizează un set de măsurări pentru partea de diagnostic, care să reprezinte caracteristicile complexului PQRST dominant. Este posibil ca parametrii finali să fie calculați pentru mai multe cicluri dominante şi să se medieze parametrii sau ca din ciclurile dominante să se alcătuiască un ciclu reprezentativ din care să se extragă parametrii.

În continuarea prelucrarii, se definesc atributele ce vor fi extrase pentru prelucrarea prin metoda recunoaşterii *pattern*-ului (va fi descrisă în paragraful privind analiza semnalului EEG).

Prevalența bolilor cardio-vasculare explică eforturile considerabile depuse pentru elaborarea programelor care au ajuns deja la un grad înalt de fiabilitate. Din cercetările efectuate până în prezent, se desprind următoarele concluzii:

- algoritmii sunt mai performanți în VCG decât în ECG standard cu 12 derivații; s-ar îmbunătăți performanțele în ECG dacă s-ar modifica derivațiile de înregistrare, astfel încât să se obțină 4 grupări de câte 3 derivații ortogonale
- ar fi necesare îmbunătățiri la detecția artefactelor, în special a *spike*-urilor, care introduc detecții fals pozitive
- în partea de *pattern-recognition* ar fi necesare îmbunătățiri ale modulelor de diagnostic, însă aceasta nu ține de algoritmii de prelucrare ci de criteriile general acceptate de comunitatea medicală, a cărei orientare spre utilizarea tehnicii de calcul va imprima reconsiderarea unor definiții, clasificări şi abordări.

1.5. ANALIZA SEMNALELOR NEPERIODICE. PRELUCRAREA EEG

Semnalele neperiodice reprezintă cel mai des întâlnit tip de semnal biologic, iar metodele de prelucrare au un caracter general. De aceea, deşi vom alege semnalul EEG ca exemplu tipic de semanal neperiodic, multe din metodele de prelucrare descrise în continuare se pot aplica şi altor semnale.

1.5.1. Caracterele generale ale semnalului EEG

Semnalul electroencefalografic reprezintă activitatea electrică a creierului şi este înregistrat în diferite poziții ale electrozilor de culegere pe scalp. Amplitudinea variază între

10 şi 200 μV iar frecvenţa între 0,5 şi 30 Hz. Neperiodicitatea este o caracteristică vizibilă şi pentru un ochi neavizat.

Originea semnalului este încă o problemă neelucidată. Deşi a existat şi opinia că rolul preponderent l-ar avea regiunile profunde ale creierului, actualmente se consideră că electrozii culeg în special activitatea scoarţei, semnalul pe fiecare electrod reprezentând o sumare ponderată (dependentă de distanţa la electrod şi de mediile intermediare) a activităţilor unei regiuni relativ întinse. Experienţele efectuate prin deplasarea electrozilor în regiuni învecitate au dus la concluzia că nu ar fi posibilă o delimitare externă precisă a diferitelor regiuni. Pentru a putea compara diferite trasee, s-au standardizat unele variante de amplasare a electrozilor, precum şi condiţiile tehnice de înregistrare. (Din păcate există încă mai multe standarde).

Activitatea EEG înregistrată cuprinde mai multe elemente cu diferenţe regionale şi cu evoluţie dinamică, ceea ce ridică numeroase dificultăţi de prelucrare, dar totodată incită curiozitatea ştiinţifică de a găsi originea acestor variaţii, cu scopul ca apoi să poată fi utilizate (de exemplu, în scop diagnostic) sau chiar controlate (terapie). Iată o clasificare a activităţilor EEG efectuată de Dumermuth:

a) Activitate spontană neparoxistică:

- activităţi fără modificări temporale semnificative (alpha, beta, ritm lent continuu, activitate lentă polimorfă)
- activităţi cu modificări lente în timp (activitate în somn, activitate postictală, activitate fluctuantă în comă, activitate de hiperventilaţie, descărcări *seizure)*
- activităţi de tip intermitent (activitate sigma în formă de *spindles* de somn, ritmul miu, ritmuri lente intermitente, diferite *pattern*-uri psihomotorii).

b) Activitate spontană paroxistică:

- vârfuri, unde ascuţite
- complexe vârf - undă
- formaţii ritmice (3/sec) vârf şi undă
- unde lente paroxistice
- vârfuri pozitive
- complexe SSLE
- complexe K şi potenţiale vertex în somn.

c) Activitate evocată:

- potenţiale evocate tranzitorii
- însuşirea fotică a ritmului
- activitate la deşteptare
- efectele de închidere a ochilor
- undele λ.

În mod uzual un medic de explorări funcţionale, în interpretarea pe care o dă unui traseu EEG utilizează o terminologie specifică: de exemplu "traseu iritativ (unde mai ascuţite), cu frecvente fusuri alfa, supravoltat (amplitudini mai mari) etc". Prin metodele de prelucrare se vor putea estima prin valori numerice caracteristicile uzuale adăugându-se şi o serie de parametri noi.

În ciuda faptului că eforturile depuse pentru analiza semnalului EEG acoperă o paletă largă şi variată de metode, rezultatele obţinute până în prezent sunt relativ modeste, însă acest domeniu este deosebit de dinamic apărând mereu metode noi.

1.5.2. Clasificarea metodelor de prelucrare

Complexitatea semnalului a determinat apariția unui număr mare de posibilități de abordare a prelucrării, fiind împărțite în două mari categorii:

- metode elementare de analiză
- metode integrative.

Metodele elementare de analiză pot fi divizate la rândul lor în două mari clase, după aspectul preponderent urmărit în prelucrare:

- prelucrări în domeniul timp (analize temporale);
- prelucrări în domeniul frecvență (analize frecvențiale).

1.5.3. Metode elementare de analiză. Analiza temporală a semnalului EEG

Analizele temporale cuprind tehnici de prelucrare care presupun ca element fundamental secvența temporală a datelor, orientate pe câte o caracteristică particulară a semnalului. În exemplele care urmează în acest paragraf ne vom referi, în general (cu excepția cazurilor ce vor fi menționate), la semnalul prezentat în fig. III.14, înregistrat în derivația centro-occipitală stângă, pe un subiect sănătos.

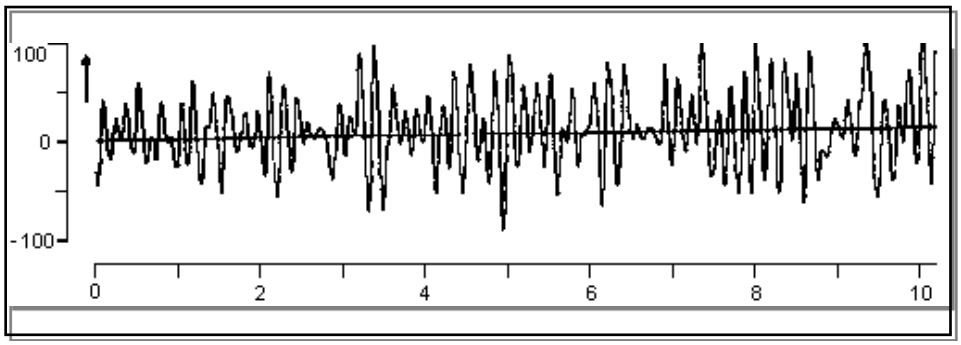


Figura III.14. Porțiune de traseu EEG; imagine realizată de programul de vizualizare. [Popescu 1988]

a) Analiza amplitudinilor (Drohocki) constă în estimarea unei funcții de distribuție a amplitudinilor şi parametrilor statistici asociați ei. Este utilă pentru rezumarea datelor EEG pe perioade lungi şi pentru caracterizarea activității spontane, astfel încât să poată fi detectate şi evenimente paroxistice.

Principala formă de prezentare a rezultatelor analizei amplitudinilor este histograma amplitudinilor (fig. III.15a). O distribuție mai ascuțită decât cea normală ar caracteriza un semnal "subvoltat" (când majoritatea valorilor sunt mici, în apropierea liniei de zero) în timp ce o distribuție mai turtită ar corespunde unui traseu ce uzual ar fi fost numit "supravoltat".

b) *Analiza intervalelor (perioadelor)* (Saltzberg şi Burch), ce reprezintă un studiu al distribuției intervalelor între "punctele specifice", cum ar fi: traversarea axei *(zero crossing)*, extreme, puncte de inflexiune, etc (fig. III.15b). Se evaluează perioade (intervale) şi pe prima şi a doua derivată a semnalului. Deşi metoda este simplă, rezultatele s-au dovedit a fi în foarte bună concordanță cu cele obținute prin metode mult mai sofisticate atunci când în semnal este prezent un ritm dominant.

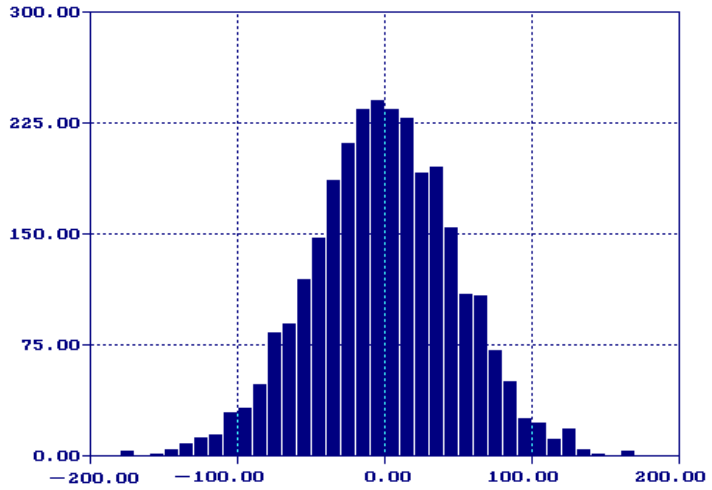


Figura III.15a. Histograma amplitudinilor, împărţind întreaga plajă de valori de intrare a semnalului (-200μV, + 200μV) în 40 clase de valori (a câte 10 μV fiecare)

Când sunt prezente mai multe frecvenţe, metoda poate conduce la interpretări eronate. Frecvenţele joase influenţează puternic linia de zero, ceea ce determină erori în special ale frecvenţelor înalte. Adăugând însă şi analiza primelor două derivate, se înlătură unele erori, dar frecvenţele joase tot nu sunt detectate suficient de bine.

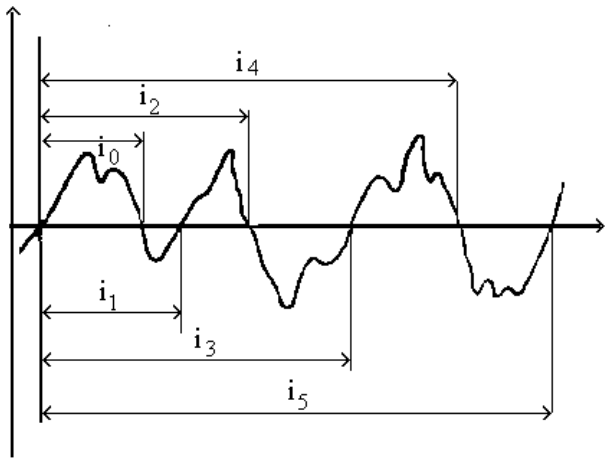


Figura III.15b. Intervalele de traversare a axei de către semnal; pentru semiunde se utilizează intervalele cu indici pari (i_0, i_2, ...), pentru unde întregi, intervalele cu indici impari (i_1, i_2, ...)

Analiza intervalelor, prin simplitatea sa şi viteza mare de calcul este potrivită pentru studii multicanal sau prelucrări de lungă durată. A fost utilizată pentru analizele EEG în timpul somnului, pentru monitorizări şi în psihofarmacologie pentru a determina profilul modificărilor induse de diferite medicamente.

c) *Analiza intervale amplitudini* (Marko şi Petsche), s-a dezvoltat în mai multe variante. Cea mai simplă este o analiză a intervalelor, la care se adaugă şi informaţii privind amplitudinea semnalului, utilizate pentru monitorizare în timpul anesteziei şi operaţiilor (Pronk 1975) şi pentru selecţia automată a epocilor fără artefacte de înregistrare (Matousek).

Cea mai uzuală variantă, numită şi *analiza secvențială* (Demetrescu), Harner şi Osterngren constă în măsurarea lungimii de undă şi amplitudinea de la vârf la vârf a fiecărei unde definite prin trecerile axei.

O altă variantă întâlnită mai frecvent şi numită *detecția vârfurilor* măsoară amplitudinea de la vârf la vârf a fiecărei unde şi perioadele între amplitudinile extreme (de fapt traversări ale axei pentru prima derivată a semnalului). Se apreciază că această metodă este cea mai apropiată de modul în care un EEG-ist citeşte o electroencefalogramă. Metoda e mai robustă în cazul variațiilor liniei de zero, însă foarte sensibilă la zgomot de frecvență mare, care introduce numeroase oscilații mărunte ale undei; se impune în aceste analize o bună filtrare a frecvențelor mari şi o "netezire" a semnalului înainte de prelucrare.

Se observă că aceste analize au ca problemă majoră definirea exactă a unei unde sau semiunde, unele variante fiind foarte sensibile la frecvențe joase, altele la frecvențe înalte. Au fost propuse şi variante combinate, utilizându-se analiza intervalelor pentru unde cu amplitudini mari şi frecvențe joase, împreună cu detectarea vârfurilor pentru undele cu amplitudini mici şi frecvențe mari (Lim şi Winters). Tot o formă a detecției vârfurilor este şi detecția înfăşurătoarelor semnalului (Schenk), linia zero fiind considerată media între înfăşurătoarea inferioară şi cea superioară. Un aspect interesant al acestor metode îl constituie posibilitatea reprezentării grafice a fiecărei unde ca un punct în sistemul de coordonate amplitudine-interval (lungime de undă), fiind posibilă şi medierea în timp real pe mai multe benzi de frecvență şi pe mai multe canale deodată, obținându-se hărți topografice ale activității EEG.

d) Analiza corelației (Barlow şi Brazier) compară un tronson (o epocă) a semnalului cu un alt tronson, fie al aceluiaşi semnal (autocorelație), fie al unui semnal cules pe un alt canal (intercorelație). Compararea se realizează prin deplasarea în timp a tronsonului comparat față de cel de referință şi efectuarea produselor tuturor perechilor de valori. Se obține, astfel, o funcție (în raport cu "deplasarea în timp") ce evidențiază componentele periodice ale unui semnal (sau două semnale).

Funcția de intercorelație, prin capacitatea sa de a detecta decaljul cu care apar unele unde în anumite regiuni, este utilă în determinarea originii anumitor unde EEG, în special a unor focare de epilepsie, care pot fi astfel localizate destul de precis.

Funcția de autocorelație poate fi utilizată pentru obținerea transformatei Fourier a semnalului, care reprezintă deja o prelucrare frecvențială.

În figura III.16 este prezentată funcția de autocorelație a semnalului luat ca exemplu în fig. III.14.

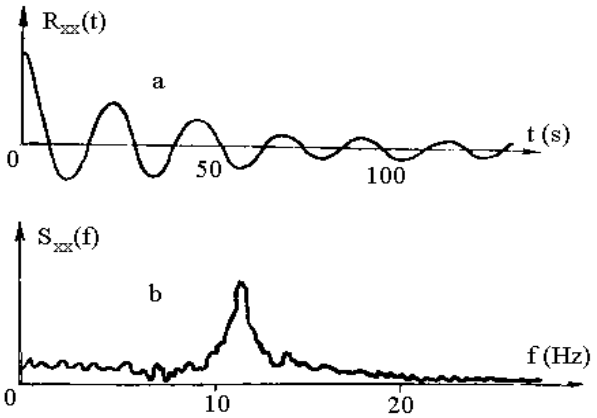


Figura III.16. Funcția de autocorelație a semnalului EEG. [Popescu 1988]

Având o foarte solidă fundamentare matematică şi fiind şi strâns legată de analiza spectrală, analiza corelației este utilizată ca un instrument de referință pentru celelalte metode. Există sisteme analogice (corelatoare) care ralizează automat analiza corelației (atât auto- cât şi intercorelația).

1.5.4. Analiza spectrală

Punctul de plecare al analizei spectrale este ipoteza că un semnal periodic complex poate fi considerat ca suprapunere a unor semnale periodice simple (de exemplu, sinusoidale), cu diferite frecvențe, amplitudini şi faze inițiale. Se obține astfel din funcția semnal o funcție spectru reprezentând o distribuție în domeniul frecvență *s(f).* Ipoteza s-a generalizat şi pentru semnale neperiodice, obținându-se, în general, un spectru continuu al semnalului. În funcție de mărimea reprezentată în spectru se disting *spectre de amplitudine* sau *spectre de putere* (în care intervine pătratul funcției semnal).

a) Considerații teoretice. Există mai multe posibilități de alegere a funcțiilor periodice simple după care să se facă descompunerea, însă cea mai uzuală descompunere este *dezvoltarea în serie Fourier,* care consideră că semnalul ar rezulta prin suprapunerea unor sinusoide

($k=1,2, ...$) cu diferite amplitudini *(M_k),* frecvențe *(f_k)* şi faze *(φ_k)* :

$$x(t) = M_0 + \sum_{k=1}^{\infty} M_k \cos\left(2 \cdot \pi \cdot f_k \cdot t - \varphi_k\right) \qquad \text{(III.5)}$$

Pentru un semnal sinusoidal spectrul este o linie, pentru un semnal periodic oarecare este un grup de linii. Deoarece analiza cuprinde întotdeauna o porțiune finită de semnal, liniile din spectrul calculat nu mai sunt foarte înguste, lățimea liniilor fiind dependentă atât de durata înregistrării, cât şi de perioada de eşantionare a semnalului. În cazul în care semnalul este neperiodic, spectrul este continuu, iar suma din relația (III.5) devine integrală. În fig. III.16 au fost prezentate câteva tipuri de semnale împreună cu spectrele lor.

Încă de la începuturile electroencefalografiei, în semnalele culese au fost distinse diferite ritmuri fundamentale (delta: 0,5 - 3 Hz, theta: 3 - 7 Hz, alfa: 8 - 12 Hz, beta: 16 - 22 Hz), însă datorită suprapunerii lor este dificil a se aprecia din ochi ponderea fiecărui ritm. Aceasta explică şi interesul deosebit acordat acestui tip de analiză, cu o solidă fundamentare matematică (analiza spectrală se utilizează şi în tehnică în studiul vibrațiilor, precum şi în electronică).

Prin analiza spectrală se calculează componentele spectrale, adică ponderea pe care o au diferite frecvențe care prin suprapunere ar genera un semnal similar cu cel analizat. Graficul care se obține se numeşte spectru, mărimea reprezentativă de obicei fiind "densitatea spectrală de putere".

b) Alegerea parametrilor de prelucrare a semnalului real

În toate cazurile reale semnalul este de durată limitată ΔT (de ordinul secundelor); *durata tronsonului preluat* se mai numeşte şi "epocă". Acest interval ΔT determină *rezoluția spectrală* Δf, adică distanța minimă între două linii spectrale ("finețea" de reprezentare a spectrului care se întinde între 0 şi f_{max} (F_{NY} - recvența Nyquist). Relația între Δf şi ΔT este:

$$\Delta f = 1/\Delta T \qquad \text{(III.6)}$$

Conform acestei relații observăm imediat că, pentru a obține o rezoluție spectrală satisfăcătoare (preferabil sub 1Hz), avem nevoie de epoci destul de lungi, ceea ce înseamnă achiziționarea unui număr ridicat de puncte, deci şi creşterea duratei de prelucrare.

Frecvența de eşantionare f_e şi durata ΔT a unei epoci prelucrate (lungimea tronsonului) determină structura spectrului obținut. Deoarece semnalul EEG spontan nu interesează, în general, frecvențele mai mari de circa 30 Hz (F_{Ny}), rezultă că frecvența de eşantionare trebuie să fie $f_e \geq 60$ Hz. Însă, în aceste situații, semnalul trebuie să fie bine *filtrat,* înlăturându-se toate frecvențele $f > f_{max}$ în caz contrar apare fenomenul numit *aliasing,* adică frecvențele superioare sunt interpretate ca frecvențe joase şi nu există nici un procedeu de a înlătura acest fenomen după eşantionare. În cazul în care semnalul nu este bine filtrat, este preferabil să se lucreze cu frecvențe de eşantionare mai ridicate (în dauna timpului de rulare).

În cazul unui semnal aleator, variabilitatea spectrului nu scade la creşterea duratei înregistrării, astfel încât pentru mărirea reproductibilității se utilizează neteziri ale spectrului sau chiar medieri ale spectrelor obținute pe perioade succesive mai scurte. Trebuie remarcat că spectrul obținut printr-o transformare Fourier pe un tronson mai lung nu este identic cu cel obținut prin medierea spectrelor pe mai multe epoci succesive mai scurte ce acoperă acelaşi tronson (cu cât componenta periodică a semnalului este mai ridicată, cu atat diferențele sunt mai mici). Un alt motiv al variabilității este însaşi natura semnalului, a cărei evoluție în timp este reprezentată şi prin evoluția spectrului. Ar fi însă necesar ca spectrul, prelucrat prin metodele de mai sus, să nu difere prea mult în aceleaşi condiții experimentale. Acest deziderat este atins dacă semnalul cules respectă anumite condiții.

Cu relațiile prezente mai sus putem calcula parametrii de prelucrare ai unui semnal.

Exemplu. Înregistrăm un semnal EMG cu valori în plaja 0-10 μV, utilizând un CAN pe 8 biți, cu frecvența de eşantionare de 500 Hz, preluând epoci de câte 2 secunde. Să estimăm:

a) perioada de eşantionare (în ms)
b) frecvența maximă în spectru
c) numărul treptelor de cuantizare
d) valoarea unei cuante de amplitudine (ce variație de potențial corespunde unui bit)
e) rezoluția spectrală.

Deci datele problemi sunt: Vmin = 0, Vmax = 10μV, n = 8 biți; fe = 500 Hz, ΔT = 2s.

Rezolvare.

a) Conform relației III.1:
   $Te = 1/fe = 1/500 = 0.002\ s = 2\ ms$
b) Din teorema de eşantionare III.2:
   $F_{max} = F_{NY} = fe/2 = 250\ Hz$
c) Numărul treptelor de amplitudine este conform III.3:
   $N = 2^n = 2^8 = 256$
d) O treaptă de amplitudine (cuantă), confrm III.4 are valoarea:
   $\Delta V = (V_{max} - V_{min}) / N = (10 - 0) / 256 \cong 0{,}04 \mu V.$
e) Rezoluția spectrală este conform III.6:
   $\Delta f = 1/2 = 0{,}5\ Hz$

Deci vom obține un spectru pentru frecvențe din 0.5 în 0.5 Hz de la 0 la 250 Hz.

c) Teste pentru semnalul EEG

O caracteristică fundamentală cerută pentru prelucrările semnalului este *staționaritatea,* adică menținerea compoziției diferitelor frecvențe aproximativ constantă. Deşi s-au dezvoltat metode ce dau rezultate confidente şi în cazul semnalelor nestaționare, problema rămâne totuşi deschisă şi în privința alegerii lungimii tronsonului; lucrând cu epoci scurte ($T \leq 1s$), vor fi sesizate toate nestaționaritățile (evenimentele tranzitorii vor modifica spectrul), în timp ce epocile mai lungi (sau medierea spectrelor pe tronsoane succesive însumând o epocă mai lungă), reduc nestaționaritățile şi cresc reproductibilitatea; în aceste cazuri se lucrează pe epoci de 4 - 5s, uneori şi 10s.

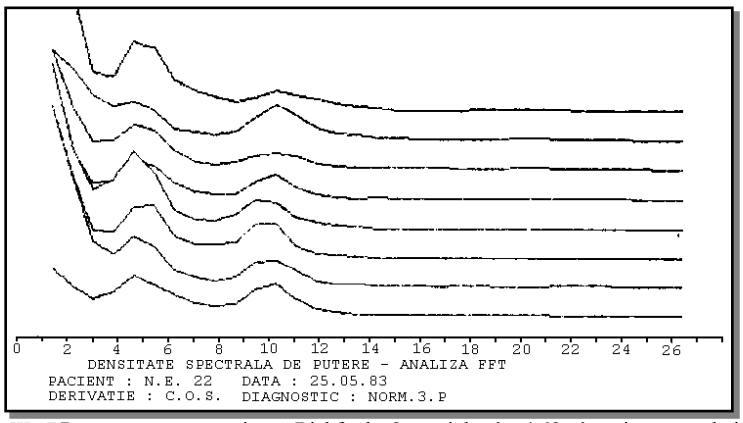
O altă caracteristică este *ergodicitatea;* mediile temporale sunt egale cu mediile statistice. Validitatea acestei ipoteze ne permite să înlocuim medierea pe mai multe realizări (pe ansamblu) cu medierea temporală (pe o singură realizare).

Se mai întâlneşte frecvent şi ipoteza *normalității* semnalului, condiție considerată uneori ca prea restrictivă şi nesatisfacerea ei nu ar determina eliminarea epocii respective din studiu.

Un test obligatoriu însă este *testul tendință,* care verifică păstrarea condițiilor de înregistrare; în anumite situații este posibilă o "deplasare"a liniei de zero a semnalului, ceea ce modifică substanțial rezultatele unor prelucrări (ex. *zero-crossing)*. Există posibilitatea ca această "alunecare" a liniei zero să fie compensată prin programul de calculator.

d) Rezultatele analizei spectrale

Spectrele obținute prin analiza Fourier a unui semnal divizat într-o succesiune de epoci se reprezintă, în mod obişnuit, într-o formă comprimată (Bickford), creându-se impresia unei reprezentări tridimensionale, deci în afară de axele pentru spectrul *S(f)* sau *G(f)* se adaugă axa timp; fiecare spectru este desenat în spatele celui anterior, având originea deplasată (cel mai adesea pe verticală); regiunile care "nu se văd" din cauza spectrului din față nu se trasează (metoda folosită în acest caz se numeşte "metoda liniei de orizont", prin asemănarea imaginii obținute cu imaginile unor lanțuri muntoase). În fig. III.17. este redată o astfel de reprezentare pentru un subiect sănătos.



Fgiura III.17 Reprezentarea comprimată Bickford a 8 epoci de câte 4,68 s înregistrate pe derivația centro-occipitală stângă a unui subiect sănătos. [Popescu 1988]

Modificările în unele stări patologice sunt tipice şi evidente. În fig. III.18 este redat spectrul unui pacient cu insuficiență renală cronică, înainte de dializă. Se observă

dispariţia ritmului alfa. După dializă spectrul devine normal, degenerând însă din nou înaintea următoarei şedinţe de dializă.

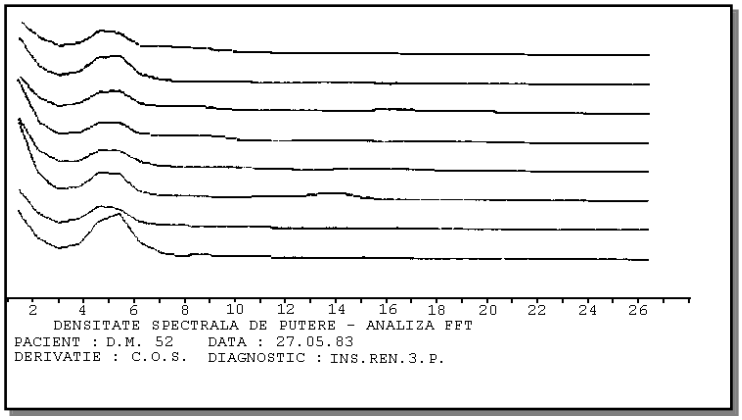


Figura III.18. Reprezentarea comprimată Bickford a 8 epoci de câte 4,68 s ale unui pacient cu insuficienţă renală cronică înainte de dializă. Puterea semnalului este mai mică în toate regiunile, iar ritmul alfa lipseşte. [Popescu 1988]

Modificările spectrelor permit şi urmărirea tratamentului efectuat în diferite tipuri de epilepsie (Gersch), diferite stări fiziologice (de exemplu diferite temperaturi - Pronk, fig. III.19), monitorizare în timpul operaţiilor pe cord, studii de psihofarmacologie (Fink), modificări cu vârsta (Turner), etc.

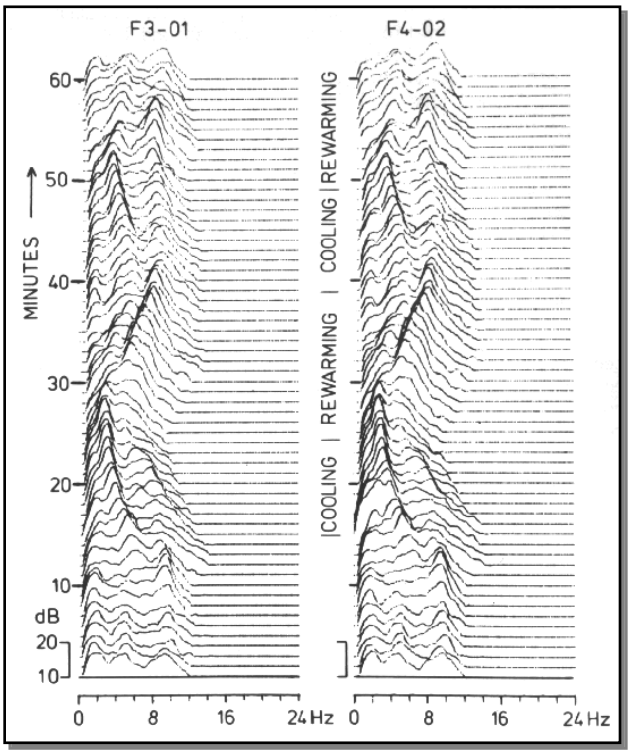


Figura III.19. Modificări spectrale induse de modificarea temperaturii. [Popescu 1988]

1.5.5. Detecţia elementelor tranzitorii

Analiza spectrală se aseamănă oarecum cu "holografia": toate regiunile (punctele) semnalului iniţial contribuie câte puţin la imaginea finală (spectru). Aceasta reprezintă un dezavantaj în cazul elementelor tranzitorii, de exemplu descărcări "spike" sau complexe "vârf-undă" care, deşi au uzual amplitudini mari, sunt de scurtă durată, deci vor contribui doar în mică măsură la spectru (fig. III.20).

O metodă simplă este folosirea analizei spectrale dar nu pe tronsoane lungi (în care contribuţia unui spike este diluată) ci pe tronsoane scurte (sub 0.5 s). Astfel contribuţia elementului tranzitoriu devine importantă şi detecţia uşoară, chiar dacă se pierde din rezoluţia spectrală. Se iau succesiv mai multe astfel de tronsoane, numite în această metodă şi "ferestre". Această metodă, propusă de Berg, se mai numeşte "metoda ferestrei mobile" şi adesea aceste ferestre se iau întreţesute pentru a nu pierde evenimente care ar putea fi eventual fragmentate în două ferestre succesive (fig. III.20)

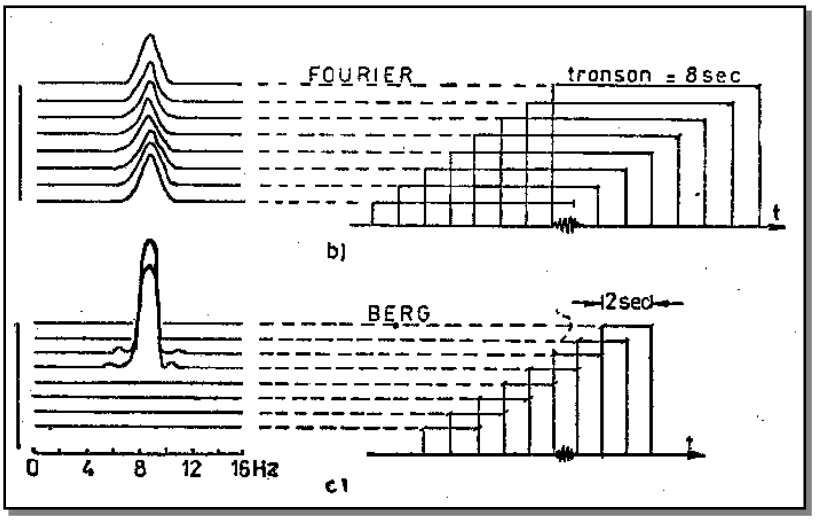
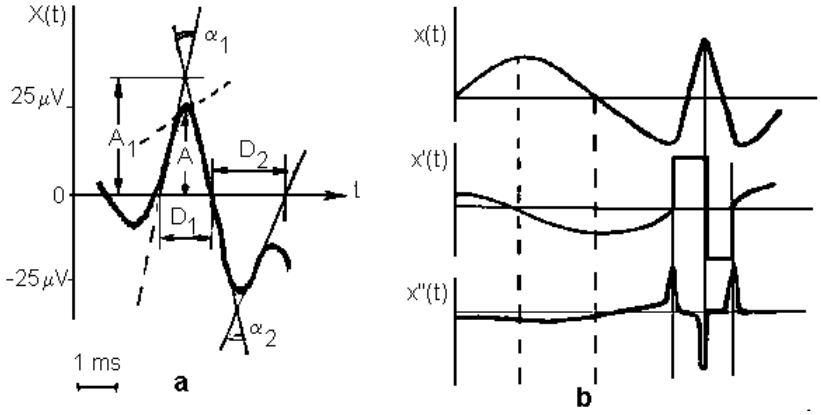


*Figura III.20.*Metoda ferestrei mobile pentru detecţia elementelor tranzitorii (Berg). [Popescu 1988]

O abordare cu totul diferită este deci necesară pentru detectarea *spike*-urilor şi undelor ascuţite (miu), elemente care sunt de mare importanţă la citirea unei electroencefalograme. Pentru detectia *spike* - urilor de către calculator (Kooi, 1996) s-au definit o serie de parametri, care sunt comparaţi cu standarde: panta ramurii ascendente, panta ramurii descendente, unghiul de vârf, amplitudinea, durata etc. (fig. III.21.a). Deoarece evenimentele din *spike*-uri sunt foarte rapide, filtrarea semnalului utilizată înainte de înregistrare diminuează frecvenţele înalte (de obicei nu le elimină de tot), împiedicând o evaluare exactă a fenomenelor rapide; chiar frecvenţa de eşantionare utilizată pentru analizele temporale sau/şi frecvenţiale ar putea fi prea mică pentru o analiză detaliată a *spike*-urilor. Teoretic, parametrii pot fi definiţi satisfăcător, dar calculul lor este laborios, fiind necesare extrapolări pentru evaluare.



Fiura. III.21. Detecția *spike*-urilor. a) Caracteristicile uni *spike;* panta nu se calculează cu punctele din vârf, care datorită eşantionării diminuează puternic panta; b) aspectul derivatelor I şi II în cazul *spike*-urilor, comparativ cu variațiile mai lente. [Popescu 1988]

O altă tehnică de detectare a *spike*-urilor este calculul derivatelor de ordin I şi II ale semnalului (fig. III.21.b), ceea ce a dus şi la realizarea unor dispozitive analogice, simple şi fidele, dedicate unor tipuri de unde. Totuşi, tehnicile numerice au o răspândire mai largă decât cele analogice.

1.5.6. Metode parametrice

Un element esențial căutat de aproape toate metoele de prelucrare este reducerea, adică sintetizarea informației conținute de semnal într-un număr cât mai unic de parametrii. Metodele enumerate mai sus prezintă toate un oarecare grad de reducere (comprimare), însă sunt adesea însoțite încă de reprezentări grafice, iar numărul parametrilor caracteristici este încă ridica. Pare deci firească tendința de căutare şi pentru semnalul EEG a unui număr redus de parametri care să conțină "toată" informația despre semnal. Au fost numeroase propuneri, unele metode au fost chiar primite cu entuziasm însă în cele din urmă s-au dovedit toate insuficiente pentru a comprima într-o manieră atât de simplă un semnal atât de comlex. Abia metodele integrative, ce vor fi expuse după acest paragraf, izbutesc să comprime satisfăcător semnalul real. Ele folosesc şi parametrii propuşi în metodele numite "parametrice". Să ne oprim la cele mai importante metode.

a) Metoda parametrilor statistici. S-au propus indicatorii statistici uzuali (medie, deviație standard, momente, asimetrie, exces), deşi corelația lor cu diferite stări fiziologice sau patologice nu a fost prea puternică.

b) Metoda descriptorilor normalizați (Hjorth) a atras de la început, prin reducerea descrierii semnalului la trei parametri numiți *activitate, mobilitate şi complexitate.*

Activitatea

$$A = \sigma_a^2 \qquad \text{(III.7..a)}$$

unde σ_a^2 este variația amplitudinilor.

Mobilitatea

$$M = \sigma_\alpha / \sigma_a \qquad \text{(III.7.b)}$$

σ_d fiind deviația standard a pantelor, iar σ_a a amplitudinilor; cum curba pantelor este de fapt prima derivată a semnalului, mobilitatea poate fi considerată o frecvență medie.

Complexitatea

$$C = \sqrt{\sigma_{dd}^2 / \sigma_d^2 - \sigma_d^2 / \sigma_a^2} \qquad \text{(III.7.c)}$$

σ_{dd} fiind deviația standard a vitezei de variație a pantei, deci este legată de derivata a doua a semnalului.

Construirea unor electroencefalografe care realizează automat şi analiza Hjorth pe fiecare din cele 16 canale (de ex. Mingograf Siemens) a determinat o răspândire mai largă a metodei. O prezentare comparativă a metodei descriptorilor Hjorth cu alte metode a fost făcută de Irwin.

În dezvoltarea acestei metode, o importantă contribuție a fost adusă de şcoala românească de neurologie reprezentată prin lucrările lui C. Arseni şi L. Popoviciu. S-au elaborat, astfel, "hărți computerizate" ale creierului pentru numeroase cazuri normale, în diferite stări fiziologice (somn) şi patologice (epilepsii, tumori), obținându-se o semiologie recunoscută pe plan mondial. Sunt deosebit de interesante evoluțiile în timp ale liniilor izopotențiale, fiind astfel pentru prima dată descrise fazele intime ale declanşării acțiunii focarelor epileptice. Compararea hărților computerizate cu imaginile obținute prin tomografie au adus noi date privind cunoaşterea fenomenelor cerebrale.

Actualmente se încearcă dezvoltarea modelelor prin completarea parametrilor şi includerea aspectelor generale în metoda recunoaşterii *pattern*-ului.

c) **Filtrarea autoregresivă Kalman** (Fenwick). Asemănarea semnalului EEG cu un semnal aleator cu distribuție normală şi medie zero ("zgomot alb") a sugerat posibilitatea descrierii sale prin parametrii unei rețele de filtre liniare, fără a presupune vreo relație a modelului cu activitatea de generare a semnalului EEG.

1.5.7. Metode integrative de analiză. Metoda "pattern recognition"

Metoda *pattern recognition*, tradusă uneori nepotrivit şi ca "recunoaştere a formelor", constituie o abordare mai largă, în cadrul căreia unii parametri estimați prin metodele anterioare să devină atribute ale *pattern*-ului. Aplicabilitatea metodei nu se limitează la prelucrarea EEG, ci este cu totul generală.

a) Principiile recunoaşterii pattern-ului

Capacitatea de a recunoaşte ceva este o caracteristică generală a ființelor umane, chiar şi a altor ființe. Procesul de recunoaştere îl practicăm în permanență: recunoaştem obiecte, melodii, tablouri, un scris de mână, paşii unui cunoscut pe scară, chiar unele stări sufleteşti. Cum? Printr-o deosebită capacitate de prelucrare a unor informații, adică un sistem complex de recunoaştere a unor atribute ale obiectului. În prima fază a procesului de recunoaştere, selectăm cele mai caracteristice atribute, obținând un rezumat tipic al informațiilor, numit *pattern*. În faza următoare, asociem acestui *pattern* un nume, comparând *pattern*-ul sesizat cu un set întreg de *pattern*-uri din memorie şi selectându-l pe cel mai apropiat.

Există două abordări principale de recunoaştere a *pattern*-ului: *metoda clasificării*, când clasele de *pattern*-uri sunt cunoscute dinainte şi *metoda grupării (clustering approach)*, când scopul este crearea şi definirea claselor.

Un sistem de recunoaştere a formelor cuprinde trei tipuri de prelucrări de date:

- achiziția datelor;
- extragerea atributelor *(feature extraction)*;
- clasificarea.

Primele două tipuri de prelucrări se efectuează utilizând cunoştinţe anterioare despre obiectele de clasificat. O cerință esențială pentru a obține rezultate bune este ca setul inițial de date să fie reprezentativ. Aceste date inițiale se împart în două categorii: *setul de "învățare" şi setul de "testare"*. Pe baza setului de învățare se determină o *regulă de decizie*, pentru a distinge *pettern*-uri din diferite clase. Calitatea clasificării pe care o poate realiza sistemul se estimează prin testarea cu datele din al doilea set.

Unii autori consideră că o recunoaştere propriu-zisă o întâlnim doar în metoda clasificării, în timp ce metoda grupării pare potrivită ca o etapă anterioară pentru definirea unor clase, când astfel de definiții nu există sau nu sunt suficient de precise.

b) Construirea unui pattern. Extragerea atributelor

Este desigur esențială faza de *extragere a atributelor* (mărimile cu capacitate de desciminare) pentru a defini un pattern, care poate fi realizată pe două căi:

i^0 *Arbordarea vectorială:* toate caracteristicile măsurate formează un vector în *spațiul măsurătorilor*; extragerea atributelor ar însemna trecerea într-un spațiu cu mai puține dimensiuni - *spațiul atributelor* (sau spațiul *pattern*-urilor), care se realizează prin definirea unui criteriu calitativ ce caracterizează distribuția *pattern*-urilor, definirea unui set de funcții pentru mapping şi selectarea funcției optime.

Definirea criteriului calitativ se face prin diverse metode asupra cărora nu ne oprim aici.

ii^0 *Abordarea structurală,* în care se consideră că rezultatul unei măsurători este o înlănțuire de elemente şi se caută elementele de bază prin care atât numărul acestora, cât şi pierderea de informații să fie minime.

c) Sinteza clasificatorului

Fiind dat un set de învățare se caută funcții de decizie ce corespund celor *K* clase. Calitatea clasificatorului se măsoară prin rata erorilor de clasificare şi de generalizare.

Există mai multe căi pentru stabilirea regulilor de decizie, cei mai uzuali clasificatori fiind cei de distanță, bazați pe reguli geometrice într-un spațiu n – dimensional.

Abordările teoretice în metoda pattern recognition sunt destul de sofisticate, iar aplicatiile de acest tip necesită resurse importante, fiind implementate mai mult pe calculatoare puternice.

d) Aplicarea recunoaşterii pattern-ului la semnalul EEG

Faptul că această metodă ia în considerare mai multe aspecte a determinat numeroşi cercetători să îmbine aspectele temporale cu cele frecvențiale, la care să adauge şi alte atribute (fig. III.22), şi să obțină astfel - o descriere mult mai fidelă a semnalului EEG. Vom reda doar câteva aspecte mai importante.

i^0 *Atribute folosite:* s-au testat numeroase caracteristici, dintre care enumerăm o parte:

- <u>atribute spectrale:</u> putere absolută sau relativă a unui ritm (δ, ϑ, α, β) frecvențele medii ale ritmurilor respective etc;

- atribute interval / amplitudine: număr de traversări axă în diferite intervale de timp, în semnal şi în prima derivată, suma amplitudinilor semiundelor etc.;

- descriptorii normalizaţi Hjorth şi coeficienţii filtrului Kalman.

ii^0 *Atribute selectate*: după ce au fost încercate pe grupe de atribute, lista s-a redus la: frecvenţa medie, puterea procentuală a undelor delta, numărul total de traversări ale axei pentru semnal şi pentru prima derivată, primii trei coeficienţi Kalman, rădăcina pătrată a activităţii, mobilitatea şi complexitatea.

iii^0 *Clasificarea,* efectuată în scopul urmăririi asimetriei semnalului, a utilizat în prealabil o ierarhizare a atributelor selectate, criteriul cel mai puternic fiind rădăcina pătrată a activităţii Hjorth.

iv^0 *Evaluarea clasificării* s-a efectuat cu setul test, luând în considerare separat erorile de clasificare fals pozitive şi fals negative, eroarea medie de clasificare fiind de circa 25%.

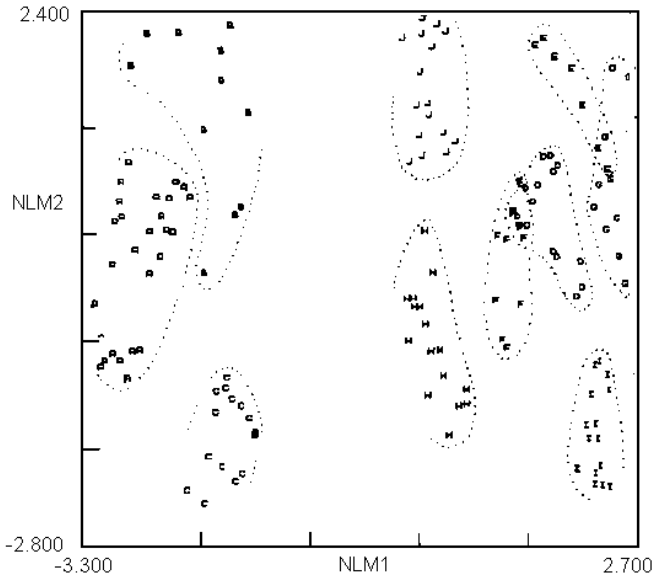


Figura III.22. Reprezentarea într-un spaţiu cu 6 dimensiuni a semnalelor EEG culese de la un pacient dializat. Grupările de puncte aparţin diferitelor perioade (înainte, în timpul şi după dializă). Prelucrare prin pachetul ISPAHAN. [Popescu 1988]

e) Avantaje şi limite ale metodei recunoaşterii pattern-ului

Această metodă, încă în curs de dezvoltare, este considerată de mare perspectivă, în special datorită faptului că ţine cont de numeroşi factori pe care îi selectează în funcţie de calitatea clasificării obţinute pe lotul de învăţare. În acest scop s-au şi creat pachete de programe specializate pentru problema recunoaşterii *pattern*-urilor (ex: ISPAHAN).

Există însă şi unele comentarii care limitează entuziasmul arătat metodei. În primul rând, este necesar un număr imens de date pentru "lotul de învăţare" şi pentru "lotul de test". Deşi nu există un criteriu concret, se apreciază că numărul de cazuri trebuie să fie de cel puţin cinci ori mai mare decât numărul de atribute considerate înmulţit cu numărul de clase. Aici ajungem în faţa unui compromis - scăderea numărului de atribute. Dar scăderea numărului de atribute considerate, ar putea influenţa calitatea clasificării (totuşi s-a arătat că "indicele de merit" este concentrat doar de câteva atribute), în timp ce creşterea

ar mări numărul de date care nu numai că ar lungi şi timpul de rulare ci, destul de des, nu avem suficiente date pentru loturile iniţiale.

Un alt dezavantaj îl constituie capacitatea de clasificare încă insuficient de ridicată (s-au realizat totuşi comparaţii cu clasificările realizate de grupuri de medici; variaţiile de clasificare, care în cazul calculatorului se numesc "erori de clasificare", au fost de acelaşi ordin de mărime).

Deşi există aceste limitări, metoda recunoaşterii *pattern*-urilor este totuşi mai complexă decât celelalte metode (incluzându-le de fapt, dacă le preia atributele).

1.5.8. Analiza sintatică

O caracteristică esenţială a semnalelor EEG este dinamica lor, în termeni tehnici fiind "nestaţionaritatea". Să ne reamintim că, din punct de vedere teoretic, aplicarea analizei frecvenţiale solicită împlinirea condiţiei de staţionaritate a semnalului. Pentru tronsoane scurte (sub 2s) această condiţie nu este îndeplinită însă ea devine satisfăcătoare de la 4-5 s în sus. Totuşi, ne putem lesne imagina că numeroase fenomene cerebrale pot fi foarte rapide şi condiţia de staţionaritate nu ar face decât să ilustreze doar o "activitate medie" a creierului care ar putea într-adevăr fi considerată aproximativ constantă pentru o anumită stare a subiectului. Analiza activităţilor cerebrale intime trebuie totuşi să acorde atenţia cuvenită evenimentelor scurte sau rapide astfel încât, cu riscul sacrificării rezoluţiei spectrale, cercetările au investigat în detaliu şi evoluţiile pe epoci scurte. Un grup de cercetare al Universităţii Vanderbilt din Tennessee a propus o metodă sintatică de analiză a semnalelor EEG, care ar cuprinde următoarele faze:

- divizarea semnalului cules (pe mai multe canale) în epoci scurte (0.3-1 s)
- efectuarea tuturor analizelor elementare pe aceste epoci
- aplicarea metodei "pattern recognition" pentru fiecare epocă; se pot astfel defini o serie de "tipuri de activităţi" care vor fi caracterizate fiecare printr-o etichetă ("label") a epocii
- succesiunea epocilor va fi reprezentată printr-o succesiune a etichetelor, care formează o "propoziţie"
- se analizează "propoziţiile" obţinute în diferite tipuri de activităţi
- se aplică analiza sintactică a acestor propoziţii pentru clasificarea activităţilor cerebrale.

Prin această metodă au putut fi clasificate câteva tipuri de activităţi cerebrale iar cercetările sunt în plin avânt, aşteptându-se ca prin îmbinarea acestor metode sintactice cu cele de "pattern recognition" să se obţină identificarea unor tipuri de stări sau activităţi ale creierului uman.

1.5.9. Potenţiale evocate

Studiul activităţii cerebrale de fond, deşi aduce o serie de date privind starea fiziologică sau patologică a subiectului, nu reprezintă decât o imagine parţială a funcţionalităţii creierului. Deşi se iau toate precauţiile pentru a nu modifica condiţiile de înregistrare, caracterul nestaţionar al semnalului apare destul de des în evidenţă. O serie de date noi apar însă atunci când urmărim ecoul la nivel cerebral al unor stimuli controlaţi de noi. Potenţialele înregistrate în aceste condiţii reflectă evident modul de reacţie a creierului la modificarea condiţiilor, contribuind la aprofundarea mecanismelor care sunt atât de puţin cunoscute. Se explică astfel şi interesul deosebit acordat studiilor de acest gen, care devin din ce în ce mai numeroase.

Natura stimulării poate fi diversă. Cele mai numeroase studii se referă la stimularea vizuală (cu lumină stroboscopică sau, mult mai des, cu modele tip tablă de şah),

stimularea auditivă (sunete de diferite durate sau înălțimi, uneori modulate cu frecvențe mai joase) sau stimularea somatică.

a) Extragerea semnalului

Problema majoră în studiul potențialelor evocate o constituie amplitudinea mică a răspunsului (câțiva μV), care fiind suprapus peste activitatea de fond, de 50 - 100 μV, în mod obişnuit nici nu se observă pe traseu. Considerând ca semnal potențialul evocat iar activitatea de fond ca zgomot, extragerea semnalului devine o problemă de creştere a raportului semnal / zgomot. Presupunând că potențialul evocat apare întotdeauna cu aceeaşi

latență după stimulare, Dawson a introdus metoda superpoziției, prin care se suprapun epoci de semnal, fiecare epocă începând în momentul stimulării. Deoarece activitatea de fond este aleatoare, regiunile fără răspuns evocat îşi compensează valorile, în timp ce regiunile ce conțin răspunsul îşi amplifică valorile (fig. III.23), apărând în felul acesta *metoda medierii,* care s-a răspândit foarte rapid după apariția mediatoarelor electronice.

O altă metodă de extragere a semnalului din zgomot este analiza Fourier într-o bandă îngustă de frecvență, în cazul în care stimulul se repetă la intervale regulate.

În studiile privind potențialele evocate se recomandă utilizarea unei frecvențe de eşantionare ridicate, deoarece fenomenele sunt rapide (sunt necesare rezoluții de ordinul milisecundei); de asemenea, fiind necesare precizii ridicate de evaluare a amplitudinilor, se recomandă convertoare de 12 - 16 biți.

b) Nestaționaritatea potențialelor evocate

Metoda medierii sau analiza Fourier ar aduce semnalul la forma sa pură dacă ar fi adevărată ipoteza că potențialul evocat ar fi identic la fiecare stimulare. Constatările experimentale arată că potențialele evocate se modifică progresiv în timp, ceea ce limitează acuratețea cu care pot fi comparate două potențiale evocate.

S-au încercat mai multe variante de a reduce erorile datorate nestaționarității, fie prin stimularea simultană a diferitelor regiuni, fie, în cazul stimulării vizuale, prin modularea intensității stimulului cu un zgomot cu repartiție normală.

Trebuie menționat şi fenomenul de *obişnuință la ritmul stimulării:* în anumite cazuri, după un set de stimuli repetați la intervale regulate, dacă se opreşte stimularea se mai obțin încă răspunsuri.

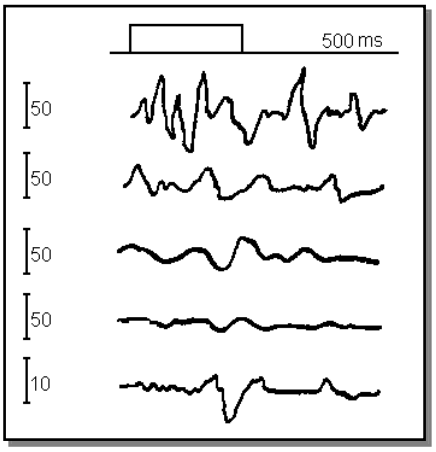


Figura III.23. Extragerea potențialului evocat prin mediere. a) evoluția stimulului luminos; b) înregistrarea unei stimulări; c) medierea a 2 stimulări; d) medierea a 5 stimulări; e) medierea a 100 de stimulări; f) semnalul din (e) amplificat. [Popescu 1988]

Pentru a înlătura efectele fenomenului de obişnuinţă, stimulatorul trebuie să genereze stimuli la intervale neregulate de timp, iar sumatorul sau mediatorul să fie de asemenea comandat de stimulator. S-au realizat dispozitive dedicate pentru aceste stimulări şi medieri, numite "*averager*".

c) Analiza potenţialelor evocate

Considerând un potenţial evocat înregistrat cu toate precauţile tehnice, putem extrage din el o serie de parametri deosebit de importanţi: latenţa, undele caracteristice (pozitive sau negative), durata lor etc. Din punct de vedere teoretic, problemele ridicate sunt similare cu cele de la analiza temporală a semnalelor (analiza intervalelor sau interval / amplitudine): definirea liniei de zero, definirea undelor etc. O serie de metode caracteristice prelucrării semnalelor de fond se aplică şi în studiul potenţialelor evocate: analiza corelaţiei, analiza Fourier, pattern recognition.

d) Clase de potenţiale evocate

S-au stabilit mai multe categorii separate de potenţiale evocate. În categorisirea unui potenţial evocat, se întâmplă destul de des ca acesta să aparţină mai multor categorii. O primă clasificare împarte potenţialele legate de procesul de *stimulare* şi potenţiale legate de *procesul de cunoaştere* - apar aşa-numitele "unde de aşteptare"; tot în această clasă intră şi undele P300 (denumirea provine de la faptul că deflexia este pozitivă şi au o durată de circa 300 ms).

O altă clasificare cuprinde categoriile: potenţiale evocate tranzitorii şi potenţiale evocate staţionare. Potenţialele tranzitorii, numite uneori şi singulare, se obţin printr-o stimulare unică. Analiza acestor semnale este mai laborioasă şi se efectuează în special în domeniul temporal. Mecanismele generării diferitelor unde şi semnificaţia lor nu sunt încă elucidate. Potenţialele staţionare se obţin prin stimularea repetată şi se analizează prin metoda medierii descrisă mai sus (fig. III.24).

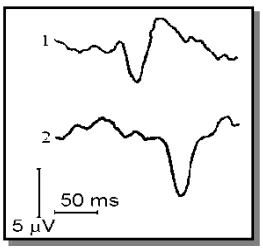


Figura III.24. Răspuns vizual evocat: (1) la un subiect normal; (2) la un subiect cu scleroză în plăci. [Popescu 1988]

Cele două metode au avantajele şi dezavantajele lor. Pentru studiul variabilelor cognitive, potenţialele evocate tranzitorii par cu mult mai bune decât cele staţionare. În acelaşi timp, potenţialele evocate staţionare pot fi prelucrate imediat de un analizator Fourier şi reprezentate ca "medie mobilă" (care se deplasează în timp). De asemenea, în studiul potenţialelor evocate este important, deseori, ca şedinţa de înregistrare să fie cât mai scurtă, metoda analizei Fourier fiind din acest punct de vedere mai potrivită decât altele.

Din cele prezente mai sus rezultă, evident, efortul mare depus pentru analiza semnalelor EEG. Deşi există şi o serie de aplicaţii clinice (diagnosticul unor epilepsii, detectarea unor tumori, monitorizarea în timpul operaţiilor etc.), majoritatea metodelor sunt dedicate cercetărilor fundamentale, care ţintesc înţelegerea mai profundă a proceselor cerebrale - ce şi cum se modifică în diferite stări fiziologice şi/sau patologice.

2. INTRODUCERE ÎN PRELUCRAREA IMAGINILOR DIGITALE

2.1. DE CE PRELUCRAREA IMAGINILOR?

Preocuparea pentru dezvoltarea unor metode de prelucrare a imaginilor vine din două arii de preocupari:

(a) imbunatatirea imaginilor pentru a fi mai uşor interpretabile de către om;

(b) procesarea datelor din imagini pentru percepţia automată ("*machine perception*") - recunoaşterea automată sau autonomă.

Tehnicile de prelucrare de imagini îşi au originea la începutul anilor 1920, cand a fost instalat cablul submarin între Londra şi New York şi primele imagini au putut fi transmise pe această cale. Timpul necesar transmiterii unei fotografii s-a redus de la ceva mai mult de o săptămână (cu vaporul) la mai puţin de trei ore. Echipamentele specializate (cântărind mai mult de 15 tone) codau imaginile înainte de transmisie şi le de-codau la recepţie, imprimandu-le. Primele sisteme *Bartlane* de codificare au avut o "finețe" de cinci nivele distincte de luminozitate (în 1921), evoluând foarte repede în următorii ani, astfel că în 1929 se ajunsese la 15 nivele.

În timpul următorilor ani metodele de procesare şi transmisie s-au dezvoltat continuu, impulsionate fiind de dezvoltarea reţelelor de televiziune. Cu toate acestea, abia în 1964 s-a utilizat tehnica de calcul la prelucrarea imaginilor: imaginile selenare transmise de *Ranger 7* aveau distorsiuni şi zgomote cu caracter regulat şi au putut fi considerabil îmbunătăţite utilizând programe de calculator. De la acest punct de pornire, tehnicile de îmbunătăţire şi restaurare a imaginilor transmise de misiunile spaţiale a devenit un lucru obişnuit: misiunile *Surveyor* pe Luna, seriile *Mariner* pe Marte, misiunile *Apollo* cu echipaj uman, etc.

Tehnicile dezvoltate au fost apoi utilizate şi în alte domenii, adaptându-se sau dezvoltându-se metode noi, specifice imaginilor prelucrate şi scopului în care ele sunt utilizate.

O sferă importantă de preocupari o constituie aplicaţiile ce au ca scop îmbunătăţirea calităţii imaginilor pentru a scoate în evidenţă conţinutul util. Ele au ca ţel final **interpretarea şi analiza făcută de catre specialiştii umani**. Astfel de aplicaţii au fost dezvoltate în medicină, geografie, meteorologie, fizică, astronomie, apărare, diverse domenii industriale. Medicina a fost întotdeauna un lider în dezvoltarea de aplicaţii datorită importanţei extraordinare a imagisticii în investigaţiile medicale.

O altă arie majoră de preocupări este **recunoaşterea automată a imaginilor** ("*machine perception*"). În acest caz efortul este concentrat pe dezvoltarea de proceduri pentru "extragerea" informaţiei imagistice într-o formă potrivită pentru prelucrarea automată şi formalizată. Printre problemele tipice care au fost în parte rezolvate şi există aplicaţii ce au depăşit deja faza de laborator (fiind utilizate în mod curent): recunoşterea automată a caracterelor (*Optical Character Readers - OCR*), sisteme de vedere artificială în domeniul industrial (linii de asamblare, controlul calităţii), prelucrarea amprentelor digitale, predicţii meteorologice, aparatură de analiză automată a probelor sanguine.

Aceste sisteme de recunoaştere automată ajung la performanţe extraordinare (viteză şi acurateţe ridicate) în aria îngustă de probleme pentru care ele sunt create. Trebuie însă accentuat că **sistemul vizual uman este neegalat în performanţă** prin varietatea mare de imagini pe care le poate prelucra şi "înţelege", ca şi prin capacitatea de adaptare la condiţii noi de percepţie în funcţie de context şi de experienţele anterioare. Acesta este unul din motivele pentru care studierea mecanismelor vizuale umane rămâne un domeniu important de preocupări nu numai pentru cercetătorii din

domeniul bio-medical, ci şi pentru cei din domeniile tehnice. Un fapt important de subliniat este că cercetarea actuală recurge, de regulă, la abordări interdisciplinare.

2.2. FUNDAMENTE. UN MODEL DE IMAGINE

Termenul de **imagine monocromă** sau simplu **imagine** se referă la o funcție a intensității luminoase, notată *f(x,y).* Ea reprezintă intensitatea (luminozitatea) *imaginii* în sensul comun al cuvântului în punctul de coordonate *(x,y)*. Cum lumina este o formă de energie, ea trebuie să fie strict pozitivă şi finită:

$$0 < f(x,y) < \infty$$

Imaginile pe care le percepem în viața de zi cu zi constau din lumina reflectată de obiectele din jur. De aceea, natura funcției *f(x,y)* poate fi caracterizată de două componente:

(1) iluminare *i(x,y)*
(2) reflectanta *r(x,y)*

$$f(x,y) = i(x,y) \cdot r(x,y)$$

Iluminarea reprezintă cantitatea de lumină incidentă în punctul respectiv şi este o caracteristică a sursei de lumină:

$$0 < i(x,y) < \infty$$

Reflectanta caracterizează proprietățile obiectului – cantitatea de lumină reflectată:

$$0 < r(x,y) < 1$$

Reflectanta este scăzută pentru obiectele pe care le percepem ca fiind negre sau închise la culoare şi are valori tot mai ridicate pentru obiectele de culoare deschisă/strălucitoare.

Pe parcursul acestei introduceri în prelucrarea imaginilor digitale, vom nota intensitatea imaginii monocrome în punctul de coordonate *(x,y)* cu *l* şi vom numi această valoare **nivel de gri** al imaginii în punctul respectiv:

$L_{min} \le l \le L_{max}$ $L_{min}=i_{min}.r_{min}$ si $L_{max}=i_{max}.r_{max}$

$[L_{min}, L_{max}]$ va constitui **scara de gri** a imaginii respective.

În practică, se obişnuieşte să se lucreze cu o scară de gri normalizată: [0 , L].
Convenția este:

l=0 este considerată a fi negru
l=L este considerată a fi alb

Valorile intermediare vor fi nuanțe de gri. Cu cat L are o valoare mai mare, cu atât finețea de prelucrare este mai ridicată. Valoarea L depinde de aplicația respectivă, de precizia cu care se lucrează, etc. Sistemul vizual uman distinge aproximativ 64 de nuanțe, dar sistemele artificiale lucrează, de regulă, cu un număr mai mare de nivele de gri.

Principiile de obținere a imaginilor nu se limitează însă la spectrul vizibil al undelor electromagnetice. Pentru imaginile obținute cu aparatele Röntgen, nivelul de gri *l* este o măsură a reflexiei sau absorbției radiației X, în timp ce în ecografie se utilizează ultrasunetele. În acest fel, se poate vizualiza/investiga interiorul corpului uman fără a-i

cauza neajunsuri majore (utilizând proceduri cât mai puțin invazive). Primul pas în procesarea imaginilor este achiziția acestora şi vom prezenta pe scurt principalele metode utilizate în imagistica medicală, deoarece imaginile obținute folosind lumina vizibilă constituie o parte infimă din imaginile medicale.

Deoarece în acest curs prelucrarea imaginilor este abordată după celelalte prelucrări de date (inclusiv semnale), considerăm oportun să prezentăm o sinteză a metodelor de procesare a datelor. Tabelul următor consideră **datele în sens general**, sistematizând metodele funcție de tipul de informație de la intrare (*input*), respectiv rezultatul obținut în urma aplicării metodei (*output*):

output / **input**	date 3-D *imagine 3-D*	date 2-D *fotografie*	date 1-D *semnal*	vector *caracteristici*	date 0-D *identitate*
date 3-D *imagine 3-D*	restaurare îmbunătățire	detecția limitelor ob.	detecția liniilor	analiza imaginilor	interpretarea imaginilor
date 2-D *fotografie*	reconstrucție	restaurare îmbunătățire	detecția limitelor ob.	analiza imaginilor	interpretarea imaginilor
date 1-D *semnal*	reconstrucție	reconstrucție	prelucrarea semnalelor	analiza semnalelor	interpretarea semnalelor
vector *caracteristici*	grafică 3-D (or. obiecte)	grafică 2-D (or. obiecte)	afişare vector	procesarea datelor	pattern recognition
date 0-D *identitate*	modelare	modelare (pict.2-D)	model/schiță (pict.1-D)	exemple	_____

Adaptată după [Van Bemmel&Musen 1997]

Subliniem faptul că această clasificare nu este totdeauna strictă şi că anumite proceduri pot fi un scop în sine, în timp ce altele sunt doar paşi intermediari pentru prelucrări mai complexe. Problemele legate de grafică şi de modelare nu vor fi abordate în această prezentare (iar cele de imagistică 3-D doar tangențial), dar considerăm sinteza utilă pentru a va putea sistematiza noțiunile de prelucrare a semnalelor privite la modul general.

2.3. NOȚIUNI ELEMENTARE DE IMAGISTICĂ MEDICALĂ

Într-o măsură foarte mare, imagistica medicală este practic imposibilă fără tehnica de calcul – ea se bazează pe principii fizice care necesită un volum considerabil de calcule pentru a reda sub forma vizuală informația culeasă. Calculatoarele se utilizează în imagistica medicală pentru:

- construirea imaginilor din măsurători ale unor parametri fizici
- re-construirea unor imagini pentru o "extragere" optimă a caracteristicilor
- prezentarea imaginilor pentru a putea fi analizate în timpul actului medical
- îmbunătățirea calității imaginilor
- măsurători efectuate pe imagini – caracteristici geometrice, de coloare/intensitate, de textură, etc.
- segmentarea imaginilor – descompunerea lor în diverse componente
- arhivarea imaginilor (stocarea şi regăsirea lor) – de cele mai multe ori implică utilizarea unor tehnici de compresie.

Multe din imaginile medicale sunt generate utilizand radiație cu lungime de undă diversă (de regulă nu în spectrul vizibil) şi sunt apoi prezentate utilizând un suport auxiliar (monitor sau film fotografic). Există însă şi imagini obținute în spectrul vizibil, ca cele obținute prin endoscopie sau cele din chirurgia estetică.

Prezentăm, pe scurt, principalele metode de obținere a imaginilor medicale.

Imagini Radiologice

Potențialul radiației X de a fi utilizată pentru investigarea nedistructivă a interiorului corpului uman a fost sesizat imediat după descoperirea acesteia de catre Julius Röentgen în 1895. Metodele au fost perfecționate continuu pentru a deveni tot mai puțin nocive şi mai precise. Lungimile de undă utilizate în domeniul medical sunt cuprinse în intervalul 0.1 — 1°A.

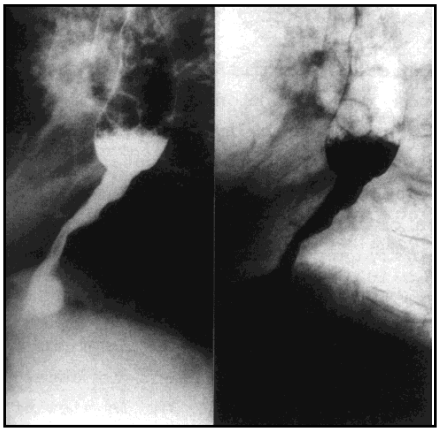


Figura III.2.3.1. Imaginea negativă (radiografia clasică - stânga) şi alternativa pozitivă (dreapta)

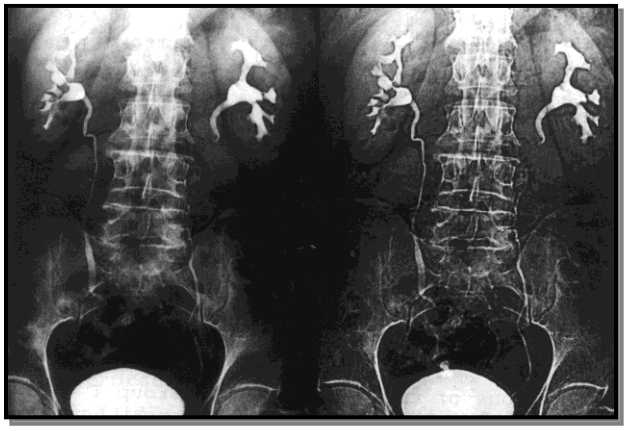


Figura III.2.3.2. Întărirea contrastului unei imagini radiografice

Radiologia clasică a utilizat filmul fotografic şi ne punem întrebarea firească dacă *este utilă prelucrarea imaginilor radiografice 2-D cu ajutorul calculatorului?* Credem că răspunsul este evident pozitiv: pe lângă facilitatea de stocare/regăsire a imaginilor, tehnica de calcul oferă posibilitatea unor prelucrări simple şi rapide atunci când este nevoie (figurile III.2.3.1 şi III.2.3.2 sunt doar nişte exemple).

O tehnică radiologică ce nu poate fi însă aplicată fără ajutorul tehnicii de calcul este angiografia digitală - **DSA *(Digital Substracted Angiography)***. Tehnica se utilizează în situații când este nevoie să se vizualizeze vase sau cavități interne ce sunt ecranate de prezența unor țesuturi cu densitate ridicată (de regulă oase). Un exemplu este cel prezentat în figura III.2.3.3, în care a fost necesară investigarea rețelei de vase

din zona craniană. S-a injectat o substanţă opacă la radiaţia X şi s-a efectuat o radiografie. Pentru zonele abdominale metoda aplicată în acest fel conduce la rezultate mulţumitoare, dar în acest caz imaginea obţinută (a) are un contrast scăzut şi, practic, în ea nu se poate distinge nimic util din punct de vedere medical. De aceea, se achiziţionează cel puţin două imagini: una înainte de injectarea substanţei de contrast şi una după. Prima imagine poartă numele de "masca" şi ea se "extrage" prin tehnici digitale din imaginea a doua, scoţându-se astfel în evidenţă tocmai diferenţa dintre cele două imagini – reţeaua de vase (b).

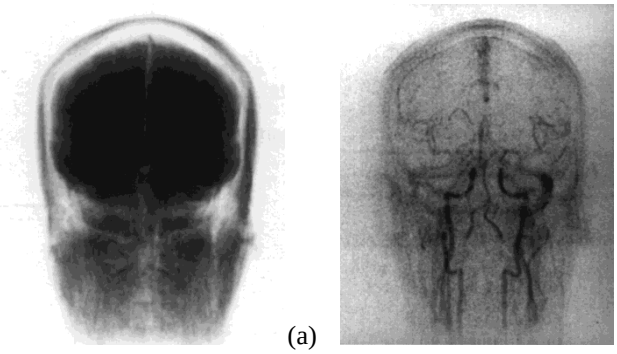

(a) (b)

Figura III.2.3.3. În *Digital Substracted Angiography* se utilizează substanţa de contrast opacă la radiaţia X, preluându-se două imagini (înainte şi după injectarea substanţei de contrast). Informaţia utilă (imaginea b în figură) se obţine "scăzând"-o pe prima (*masca*) din imaginea preluată după injectarea substanţei de contrast

Tehnica de calcul ne ajută nu numai să facem calculele necesare operaţiei de "scădere" a celor două imagini, ci şi potrivirea imaginilor prin tehnici de corelaţie (asemănătoare celor descrise la prelucrările de semnale) deoarece este practic imposibil ca pacientul să fie poziţionat identic pentru cele două imagini. Operaţia poartă numele de **"*image registration*"** şi este necesară ori de câte ori se fac operaţii de comparare, scădere, etc. a două sau mai multe imagini (ele trebuiesc translatate, rotite, scalate sau "întinse" pentru a se potrivi).

În mod normal, se achiziţionează mai multe imagini post-contrast pe măsură ce substanţa de contrast difuzează iar contrastul devine mai puternic în zonele periferice şi scade în regiunea proximală punctului de injecţie. Ele se vor combina pentru a obţine imagini de calitate în condiţiile utilizării unor cantităţi mici de substanţă de contrast şi a unei iradieri de cât mai mică intensitate.

Tomografia Computerizată

Un neajuns major al imaginilor radiologice este faptul că ele redau în două dimensiuni o realitate tridimensională – "pierderea" unei dimensiuni conduce implicit la pierdere de informaţie, lucru pe care medicii îl compensează prin cunoştinţele specifice, prin intuiţie, etc. Chiar atunci când aceste imagini sunt "îmbunătăţite" (ca în situaţia angiografiei digitale), ele rămân o proiecţie bidimensională a realităţii. Mai mult, prin tehnicile clasice este imposibil să se facă distincţie între ţesuturile moi pentru că diferenţa netă este între os şi aer, sau între substanţa de contrast şi ţesuturile moi luate împreună.

Invenţia tomografiei computerizate cu raze X (Godfrey N. Hounsfield, 1971), răsplătită cu un premiu Nobel în 1979, a produs o adevărată revoluţie în lumea medicală. Această tehnică (ilustrată în figura III.2.3.4) permite vizualizarea unor secţiuni bi-dimensionale "tăiate" în zonele de interes din corpul investigat.

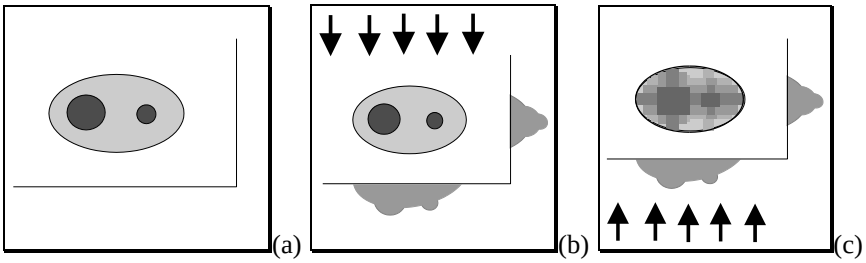


Figura III.2.3.4. Principiul proiecției inverse utilizat în tomografia computerizată. Presupunem că secțiunea pe care dorim s-o vizualizăm are forma eliptică (a) cu două formațiuni de densitate diferită în interior (de exemplu o secțiune prin antebraț). Se fac două "radiografii" (pe direcții perpendiculare) cu fascicole foarte înguste de radiație X şi se obțin *profile de absorbție* bidimensionale (b). Combinându-se cele două surse de informație, se redau diferențele de densitate în secțiunea bi-dimensională investigată (c)

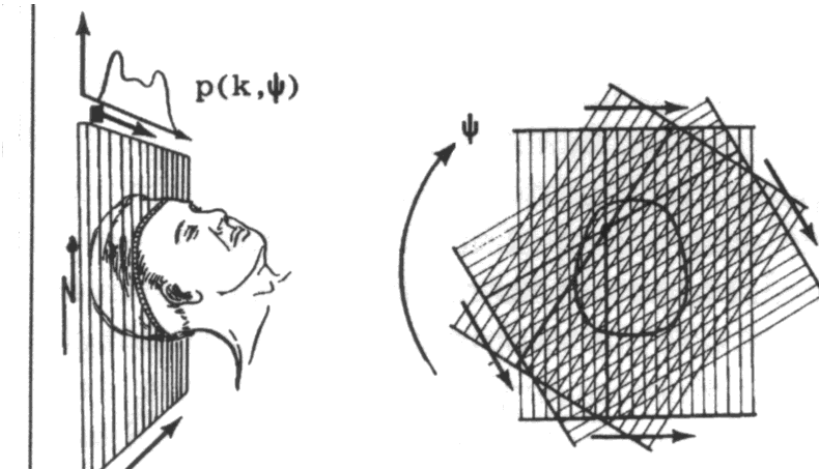


Figura III.2.3.5: Principiul tomografiei computerizate utilizat în investigațiile medicale – preluat din *[Van Bemmel&Musen 1997]*

În practică, imaginile sunt reconstruite pornind de la un număr mare de "profile de absorbție", luate la intervale unghiulare constante pentru o incidentă dată – figura III.2.3.5. Utilizând sute de astfel de profile pentru o secțiune dată, calitatea obținută este ridicată (ce a fost ilustrat în figura III.2.3.4 este o exemplificare a ceea ce s-ar obține cu doar două profile).

Tomografia computerizată are câteva neajunsuri ce derivă din faptul ca utilizează radiația X, ceea ce face uneori dificil să se vizualizeze țesuturile moi. Sunt situații în care utilizarea unei alte metode – rezonanța magnetică nucleară – conduce la rezultate mai bune. Figura III.2.3.6 ilustrează aceste diferențe.

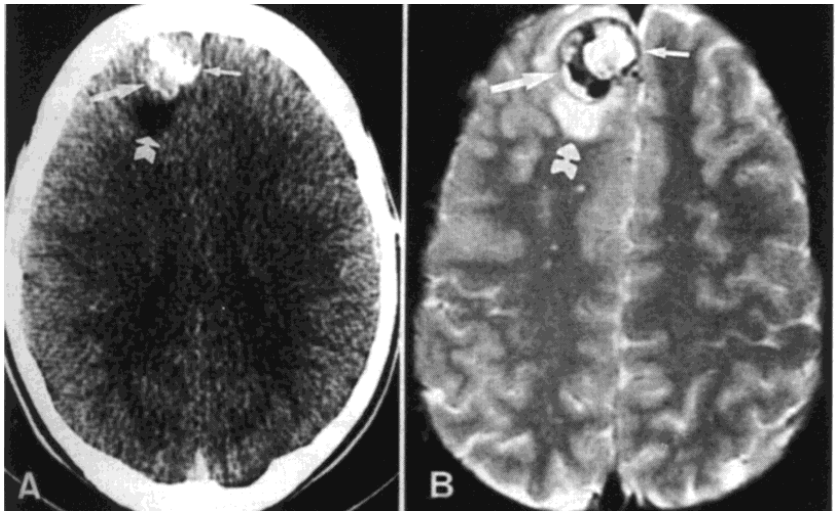


Figura III.2.3.6. Diferenţe între imagine obţinută prin tomografie computerizată cu raze X (A) şi prin rezonanţa magnetică nucleară (B)

Rezonanţa Magnetică Nucleară

Rezonanţa magnetică nucleară permite vizualizarea distribuţiei ţesuturilor din secţiuni transversale "tăiate" prin organele investigate ("felii" bi-dimensionale) – prin analogie cu tomografia computerizată cu raze X, se mai numeşte **tomografie de rezonanţă magnetică nucleară**. Numele vine de la faptul că se bazează pe un fenomen de rezonanţă între energia nucleelor atomice aflate într-un câmp magnetic şi radiaţia electromagnetică cu frecvenţă specifică fiecarui tip de atom.

Principiul este ilustrat în figura III.2.3.7 – se bazează pe momentul de spin care face ca atomii cu număr de ordine impar să se comporte ca nişte mici magneţi. Introduşi într-un câmp magneric exterior suficient de intens, ei tind să se orienteze după direcţia câmpului şi apare o magnetizare a corpului investigat. La apariţia unui impuls de radiaţie electromagnetică de frecvenţă potrivită, nucleele ce intră în rezonanţă vor absorbi energia şi vor ocupa o nouă poziţie (permisă din punct de vedere cuantic) - fenomenul se numeşte *excitaţie*.

La încetarea impulsului perturbator, nucleele vor tinde să-şi reia vechea poziţie printr-o mişcare de precesie, emiţând un semnal în radiofrecvenţa (*frecvenţa Larmor*) ce este funcţie de natura nucleelor, de combinaţiile chimice şi de condiţiile fizice în care acestea se află. Această revenire poartă numele de *relaxare* şi are (şi ea) durata în funcţie de proprietăţile ţesutului respectiv. Se măsoară doi timpi de relaxare: T_1 (relaxarea longitudinală sau *spin-lattice*) şi T_2 (relaxarea transversală sau *spin-spin*).

Utilizând succesiuni de impulsuri de excitare şi coroborând informaţiile culese (frecvenţa Larmor, T_1 şi T_2) se obţine o vizualizare a distribuţiei ţesuturilor precum şi a unor fenomene metabolice din secţiunile analizate.

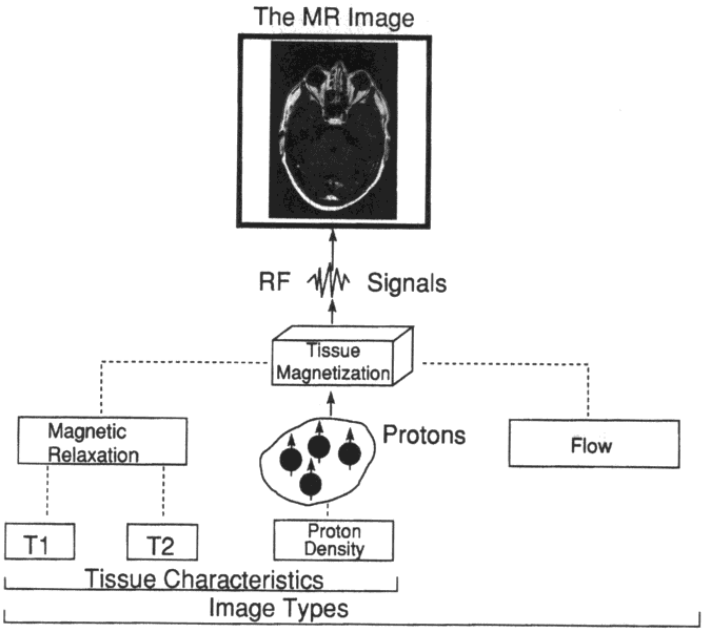


Figura III.2.3.7. Ilustrarea principiului utilizat în rezonanța magnetică nucleară pentru vizualizarea distribuției țesuturilor moi și a fenomenelor metabolice – preluat din *[Van Bemmel&Musen 1997]*

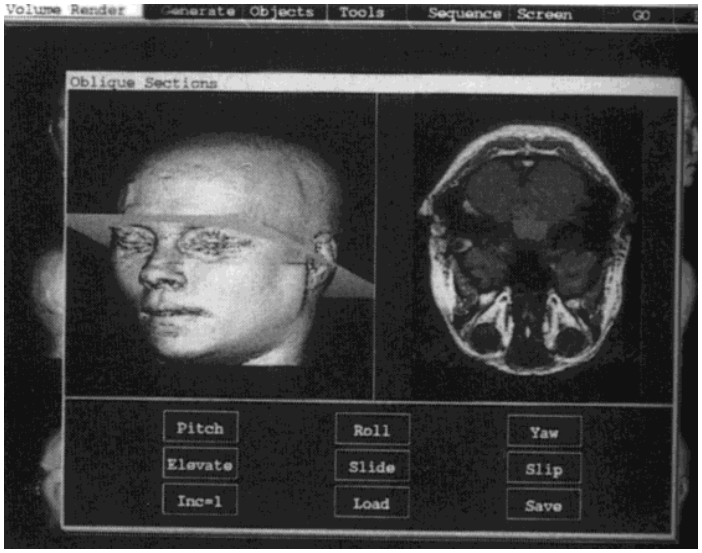


Figura III.2.3.8. Exemplu de reconstrucție a imaginilor 3-D *(3-D rendering)* într-o aplicație de imagistică medicală

Aplicațiile imagistice actuale permit reconstrucția tri-dimensională pornind de la succesiuni de imagini bi-dimensionale – figura III.2.3.8. De multe ori se combină imagini obținute prin tehnici diferite: tomografie cu raze X, rezonanță magnetică nucleară, radiografie, etc. De regulă, astfel de aplicații permit vizualizarea/reconstrucția unor noi secțiuni (ne-existente în investigația inițială) prin combinarea informațiilor culese din surse diferite.

Scintigrafia

Principiul care sta la baza acestei metode imagistice este administrarea (prin injectare sau prin inhalare) a unor substanțe marcate radioactiv (substanțe *radio-farmaceutice*) şi vizualizarea modului în care această radioactivitate "interioară" se distribuie în organism. Avantajul este că se pot obține imagini dinamice şi funcționale pentru organul investigat. De aceea, metoda este foarte utilă în investigațiile cardiace, iar sincronizarea achiziției este controlată de semnalul ECG. De mai bună acuratețe sunt dezvoltări ale metodei bazate pe principiile tomografice - SPECT (*Single Photon Emission Computed Tomography*).

O altă dezvoltare a metodei (de o acuratețe mai mare şi potrivită pentru fenomene mai rapide) a produs o adevarată revoluție în investigațiile din neuroştiințe: tomografia cu emisie de pozitroni - PET (*Positron Emission Tomography*). Ea a permis vizualizarea proprietăților dinamice ale proceselor biochimice din creier (figura III.2.3.9).

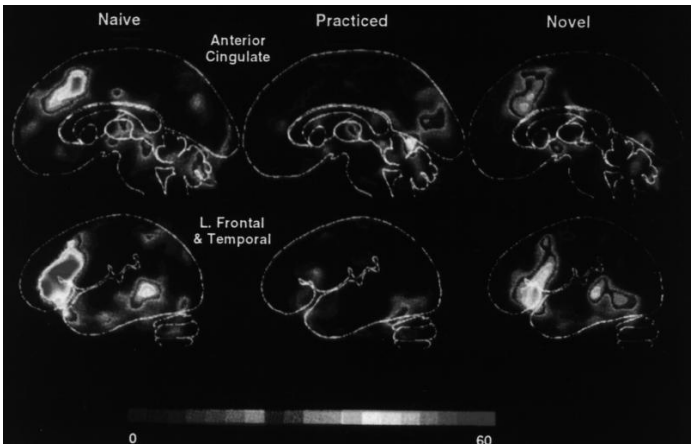


Figura III.2.3.9. Investigarea activității creierului utilizând tehnica PET – *Positron Emission Tomography.* Se introduce în circuitul sanguin deoxiglucoza care se acumulează rapid în zonele creierului care au o activitate ridicată

Ca şi în cazul celorlate metode de investigare, tehnicile scintigrafice permit reconstrucția tri-dimensională pornind de la succesiuni de imagini bi-dimensionale.

Ecografia

Tehnicile ecografice utilizează ultrasunete (unde cu frecvențe peste 20 kHz) şi se bazează pe faptul că viteza, impedanța caracteristică şi coeficientul de absorbție diferă funcție de materialul (mediul) pe care acestea îl traversează. Pentru generarea şi receptarea lor se folosesc traductori formați din cristale piezoelectrice ce pot vibra cu frecvențe cuprinse între 2 şi 10 MHz.

Datorită faptului că rezoluția şi adâncimea de penetrare a undelor impun cerințe contradictorii privind frecvența acestora, calitatea imaginilor este mai slabă decât a celor obținute prin alte metode – figura III.2.3.10.

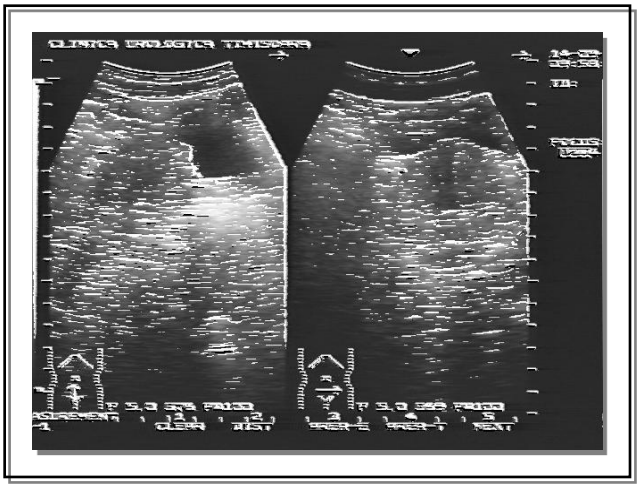


Figura III.2.3.10. Imagini ecografice ale prostatei – calitatea este inferioară celei obținute prin alte metode. Avantajul îl constituie faptul ca metoda este mult mai puțin invazivă decât radiația X sau scintigrafia

Pe lângă structurile anatomice, ultrasunetele permit şi investigarea vitezei de deplasare a unor fluide (de exemplu sângele) prin utilizarea efectului Doppler. De regulă, pentru investigațiile cardiace se utilizează combinații între imaginile anatomice şi cele dinamice.

Fiind mai puțin invazive ca alte metode (deci cu mai puține efecte secundare), investigațiile cu ultrasunete se utilizează în examinările obstetrice şi ale nou-născuților, cele oftalmologice şi cardiace, ale creierului, etc.

Termografia

Corpul uman absoarbe radiația în infra-roşu aproape fără reflexie şi, în acelaşi timp, emite radiație în infra-roşu ca o componentă a propriei energii termice. Intensitatea energiei radiante corespunde temperaturii suprafeței corpului respectiv. Pentru un subiect sănătos, temperatura corpului poate varia considerabil în timp, dar distribuția temperaturii pe suprafața pielii păstrează forme constante şi o simetrie bilaterală pronunțată.

Tehnicile termografice permit vizualizarea acestor forme ("*pattern*"-uri) şi determinarea deviațiilor de la normal şi a schimbărilor patologice. Termograful medical este constituit dintr-o camera în infra-rosu, care obține o imagine video a distributiei temperaturii pe suprafata pielii.

Termografia are câteva **avantaje** nete asupra celorlalte metode de obținere a imaginilor în domeniul medical: este complet non-invazivă şi este un sistem în timp real (schimbările pot fi sesizate cu o frecvență de o imagine pe secundă). Sigur, această tehnică nu este comparabilă cu radiografia în explorarea interiorului corpului uman, dar este complementară acesteia: radiografia furnizează informație privind structurile anatomice, în timp ce termografia indică schimbări în procesele metabolice şi de circulație.

Tehnicile termografice sunt utilizate cu succes în investigarea problemelor de circulație, reumatism, cancer de sân, localizarea placentei în timpul sarcinii, identificarea şi localizarea unor tumori intraoculare şi orbitale.

2.4. PROIECTUL *VISIBLE HUMAN*

Proiectul *Visible Human* a fost sponsorizat de U.S. National Library of Medicine cu scopul de a pune la dispoziția cercetătorilor un set de imagini de referință ale corpului uman pentru: studii de anatomie, cercetare, dezvoltarea unor aplicații pentru educație, diagnostic şi planificarea tratamentelor, simulări, realitate virtuală.

În prima fază s-au obținut un set de imagini CT (tomografie computerizată), MRI (rezonanță magnetică nucleară) şi criosecțiuni pentru:

- bărbat – rezoluție 1mm (15 Gb)
- femeie – rezoluție 0.33mm (40 Gb)

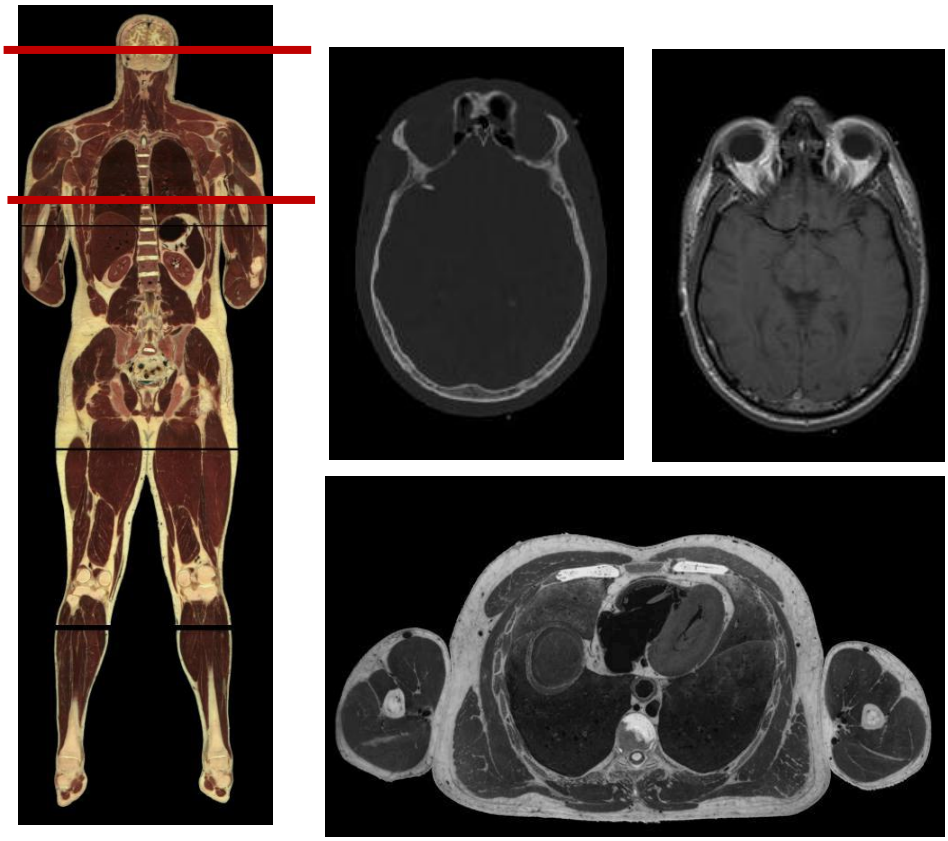


Figura III.2.4.1. Imagini realizate în cadrul proiectului *Visible Human.*

Figura III.2.4.1 ilustrează o criosecțiune longitudinală cu evidențierea unor secțiuni transversale pentru care sunt prezentate imaginile CT şi MRI, precum şi o criosecțiune la nivelul toracelui.

Mai multe informații găsiți pe site-ul dedicat acestui proiect şi dezvoltărilor sale ulterioare. Aceste dezvoltări cuprind aplicații extrem de diverse: sisteme prototip pentru screening (de exemplu în cancerul de colon), antrenarea pentru diverse proceduri chirurgicale (de exemplu cancerul de prostată, chirurgia estetică), sisteme educaționale pentru disecții anatomice, sisteme de realitate virtuală, etc.

2.5. EŞANTIONAREA ŞI CUANTIZAREA IMAGINILOR

Similar cu prelucrarea semnalelor, imaginile trebuiesc transformate într-o versiune numerică (digitală) printr-un proces de eşantionare, respectiv cuantizare. Pentru a putea fi prelucrată numeric (digital), funcția imagine *f(x,y)* va fi digitizată atât spațial, cât şi în amplitudine.

Eşantionarea şi cuantizarea uniformă

Digitizarea coordonatelor spațiale *(x,y)* este numită **eşantionarea imaginii**.
Digitizarea amplitudinii funcției *f(x,y)* este numită **cuantizarea nivelelor de gri**.

Să presupunem că o imagine continuă *f(x,y)* este aproximată prin eşantioane aranjate la intervale egale sub forma unui tablou de dimensiune N*M, unde fiecare element al tabloului reprezintă o cantitate discretă (o valoare numerică). Această matrice se va numi ***imagine digitală***:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & & f(1,M-1) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix}$$

Fiecare element al tabloului poartă numele de **element de imagine** (***picture element***) sau **pixel**. Figura III.2.5.1 prezintă intuitiv noțiunea de **pixel** şi analogul acesteia pentru spațiul tri-dimensional (**voxel – *volume element*).**

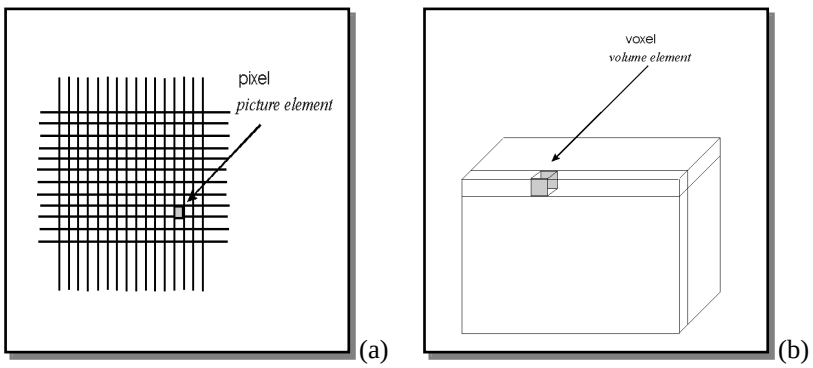


Figura III.2.5.1. Elementele unei imagini bi-dimensionale – *pixeli* (a) şi ai unei imagini tri-dimensionale – *voxeli* (b).

Exprimarea eşantionării şi cuantizării în termeni matematici formali este utilă pentru prezentările următoare.

Fie *Z* mulțimea numerelor întregi, iar *R* mulțimea numerelor reale. Procesul de eşantionare poate fi privit ca o partiționare a planului *xy* printr-o grilă în care coordonatele centrului fiecăruia din elementele grilei vor reprezenta o pereche de elemente ale produsului cartezian $Z \times Z$ (sau Z^2) - adică mulțimea tuturor perechilor ordonate de elemente *(a,b)*, cu $a,b \in Z$.

De aici, *f(x,y)* este o imagine digitală dacă *(x,y)* sunt întregi din $Z \times Z$ şi *f* este o funcție care atribuie fiecarei perechi *(x,y)* o valoare reprezentând nivelul de gri (un

număr aparţinând mulţimii numerelor reale *R*). Dăcă nivelele de gri sunt de asemenea numere întregi, funcţia *f* va avea valorile în *Z* (aşa se întâmplă de obicei).

Procesul de digitizare necesită stabilirea valorilor pentru fiecare pixel. În procesarea imaginilor digitale se obişnuieşte ca aceste valori să fie puteri ale lui 2:

$$N = 2^n \qquad M = 2^k$$

şi

$$G = 2^m$$

unde cu *G* s-a notat numărul nivelelor de gri.

Vom presupune că cele *G* valori atribuite pentru nivelele de gri sunt distribuite în mod egal (echidistant) între 0 şi L pe scara de gri. Numărul de biţi necesari pentru a memora o imagine digitală va fi dat de relaţia:

$$b = N * M * m$$

Deoarece *imaginea digitală* (matricea) reprezintă o aproximare a unei imagini reale continue, o întrebare se impune în acest moment: *cât trebuie să fie N şi M şi câte nivele de gri sunt necesare pentru o bună aproximare?*

Rezoluţia unei imagini (măsura în care se pot discerne detaliile) depinde puternic de aceşti parametri - cu cât valorile alese pentru *N(M)* şi *m* sunt mai mari, cu atât imaginea digitală se va apropia mai mult de imaginea originală.

Este dificil să definim o "imagine bună", deoarece calitatea percepută pentru o imagine este puternic subiectivă şi depinde în mare masură de cerinţele aplicaţiei care procesează imaginea. Figura III.2.5.2 prezintă imagini cu rezoluţie diferită.

Imaginea din figura III.2.5.2(a) are o rezoluţie de 287*260 pixeli cu 256 ($=2^8$) nuanţe de gri, pentru ca în figura III.2.5.2(b) rezoluţia spaţială pe fiecare axă să scadă la jumătate şi apoi la o pătrime din rezoluţia iniţială (c). Numărul nivelelor de gri a rămas constant în (a), (b) şi (c). Se poate observa creşterea progresivă a granulaţiei şi înrautăţirea calităţii în delimitarea muchiilor. Sigur că aceste detalii depind în mare măsură şi de performanţele echipamentului de tipărire, precum şi de mărimea relativă a obiectelor din imagine.

Plecând tot de la imaginea din figura III.2.5.2(a), puteţi vedea efectul scăderii numărului de nivele de gri la 16($=2^4$) în (d); numărul acestora scade apoi la 2($=2^1$) nivele în imaginea (e), care este o imagine binară (în alb şi negru). În (d) şi (e) rezoluţia spaţială a fost menţinută constantă, egală cu cea din imaginea iniţială (a). Observaţi fenomenul de "falsă conturare" încă din figura III.2.5.2(c), datorat insuficienţei nivelelor de gri utilizate pentru reprezentare.

Rezultatele prezentate arată efectele produse asupra calităţii imaginii de variaţia lui *N(M)* şi *m* luate independent. Totuşi, aceste rezultate răspund doar parţial întrebării, pentru că trebuie luată în considerare şi relaţia dintre cei doi parametri, relaţie exprimată de aşa-numitele **curbe de izo-preferinţă** (corespunzătoare imaginilor de egală calitate subiectivă).

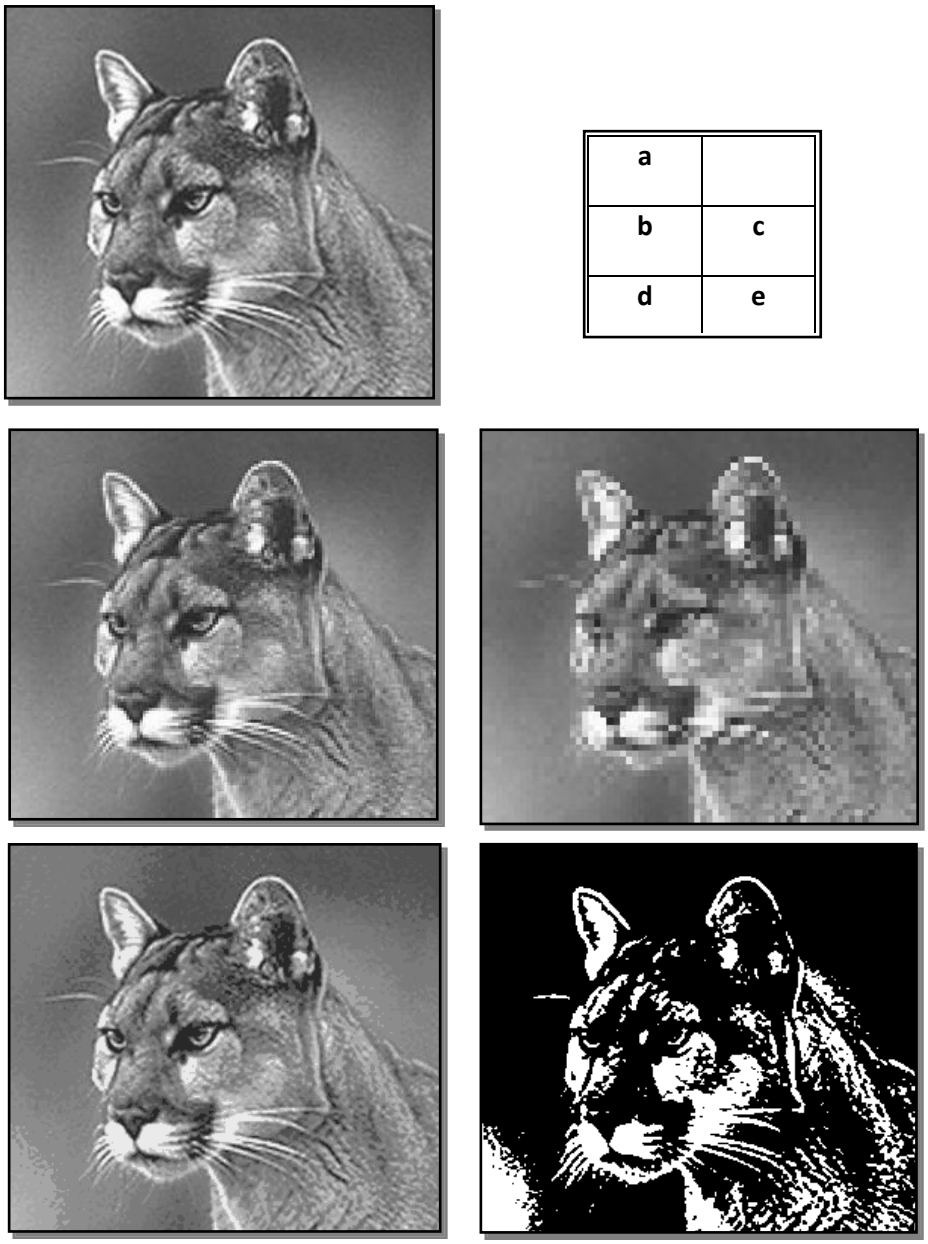


Figura III.2.5.2. Ilustrarea efectelor de schimbare a rezoluției spațiale, respectiv a numărul de nivele de gri ale unei imagini (287*260 pixeli cu 256 nuanțe de gri). Prelucrare cu *Image-Pro Plus v.3.0*

Concluziile unui studiu de izopreferință arată următoarele *[Gonzalez&Woods 1992]*:

(1) Aşa cum era de aşteptat, calitatea imaginii creşte cu creşterea lui *N(M)* şi respectiv *m*. În unele cazuri totuşi, calitatea se îmbunătăţeşte prin scăderea lui *m* - explicația este că prin scăderea lui *m* se produce o creştere a contrastului aparent din imagine.

(2) Curbele tind să devină independente de *m* pe măsură ce detaliile din imagine se înmulțesc. Aceasta sugerează faptul că pentru imagini foarte detaliate sunt necesare puține nivele de gri (rezoluția spațială este cea decisivă).

(3) Curbele de izo-preferință diferă substanțial de cele pentru care *b* (numărul de biți necesari pentru memorarea imaginii) este constant.

Pentru o iluminare constantă, sistemul vizual uman poate distinge aproximativ 64 nivele de gri iar sistemele video standard sunt adaptate la această valoare - valori exprimate pe 6 biți.

Eşantionarea şi cuantizarea neuniformă

În unele situații, pentru o rezoluție spațială fixată (număr fix de pixeli în imagine), calitatea unei imagini poate fi îmbunătățită prin utilizarea unei scheme adaptive la care procesul de eşantionare depinde de caracteristicile imaginii.

În general, o eşantionare mai fină este necesară în vecinatatea tranzițiilor bruşte între nivele de gri, în timp ce o eşantionare mai grosieră poate fi utilizată în zonele relativ uniforme. Să considerăm, de exemplu, o imagine ce constă dintr-o fața umană pe un fond uniform. În mod evident, fondul nu poartă multă informație şi el poate fi reprezentat printr-o eşantionare de finețe mai scazută decât fața umană. Eşantioanele (pixelii) rămase disponibile de la partea de fond pot fi utilizate pentru a obține o finețe mai ridicată în zona feței şi astfel rezultatul de ansamblu se va îmbunătăți. În distribuția eşantioanelor trebuie acordată o atenție mai mare zonelor de graniță ale tranzițiilor dintre nivelele de gri.

Necesitatea de a identifica în prealabil zonele de graniță (chiar dacă numai grosier) este în mod clar un dezavantaj al abordării esantionării neuniforme. De asemenea, această metodă nu este practică pentru imaginile ce conțin regiuni relativ mici.

Când numărul nivelelor de gri utilizate trebuie păstrat relativ redus, utilizarea unei cuantizari inegal distribuite a acestora este, de obicei, de dorit. O metodă similară celei descrise pentru eşantionarea neuniformă poate fi utilizată pentru distribuirea nivelelor de gri din imagine. Se vor utiliza puține nivele în vecinătatea granițelor şi mai multe nivele în zonele cu o variație uniformă a nuanțelor – astfel se vor reduce falsele contururi care apar în aceste zone atunci când cuantizarea nu este suficient de fină.

O tehnică atractivă în special pentru distribuirea nivelelor de gri constă în calcularea frecvenței de apariție a tuturor nivelelor de gri permise. Dacă există intervale ce conțin nivele cu apariție mai frecventă, în timp ce alte intervale sunt mai puțin "ocupate", se va putea face o cuantizare mai fină în interiorul intervalelor "ocupate" şi una mai grosieră în celelalte subintervale din [0,L].

2.6. RELAȚII DE BAZĂ DINTRE PIXELI ŞI OPERAȚII CU IMAGINI NUMERICE

după [Gonzalez&Woods 1992]

Vom nota o imagine cu *f(x,y)*. Când ne vom referi la un pixel particular, vom utiliza litere mici (*p, q*, etc). O submulțime sau subset de pixeli din *f(x,y)* va fi notată cu *S*.

Un **pixel *p*** aflat la coordonatele ***(x,y)*** are:

- 4 vecini pe orizontală şi verticală

  (x+1,y) *(x-1,y)* *(x,y+1)* *(x,y-1)*

  această mulțime se notează $N_4(p)$ şi se numeşte "**vecinatate de 4 a lui *p***"
- 4 vecini diagonali

  (x+1,y+1) *(x+1,y-1)* *(x-1,y+1)* *(x-1,y-1)*

această mulțime se notează $N_D(p)$

$N_4(p)$ și $N_D(p)$ formează un set numit $N_8(p)$ - "**vecinatate de 8 a lui *p*"**

O notație uzuală pentru vecinii lui *p* din setul $N_8(p)$ este următoarea: 0=Est, 1=NE, 2=N, 3=NW, 4=W, 5=SW, 6=S, 7=SE.

3	**2**	**1**
4	***p***	**0**
5	**6**	**7**

Conectivitate

Conectivitatea dintre pixeli este un concept important utilizat în stabilirea granițelor dintre obiecte și a componentelor regiunilor dintr-o imagine.

Pentru a stabili dacă doi pixeli sunt conectați, trebuie determinat dacă ei sunt vecini într-un anumit sens (să zicem, sunt vecini în N_4) și dacă nivelul lor de gri satisface un anumit criteriu de similaritate (să spunem, au aceeași valoare).

Considerăm *V* ca fiind mulțimea valorilor nivelelor de gri utilizate pentru a defini criteriul de similaritate. De exemplu, într-o imagine binară $V=\{1\}$ pentru conectivitatea pixelilor cu valoarea 1. În imagini cu o scară a nivelelor de gri, pentru conectivitatea pixelilor cu valori de intensitate într-un interval. De exemplu, între 32 și 64 vom considera $V=\{32, 33, \ldots, 63, 64\}$.

Considerăm 3 tipuri de **conectivitate**:

(a) **conectivitate 4**

Doi pixeli *p* și *q* cu valori în *V* sunt 4-conectați dacă *q* este în mulțimea $N_4(p)$.

(b) **conectivitate 8**

Doi pixeli *p* și *q* cu valori în *V* sunt 8-conectați dacă *q* este în mulțimea $N_8(p)$.

(c) **conectivitate m** (conectivitate mixtă)

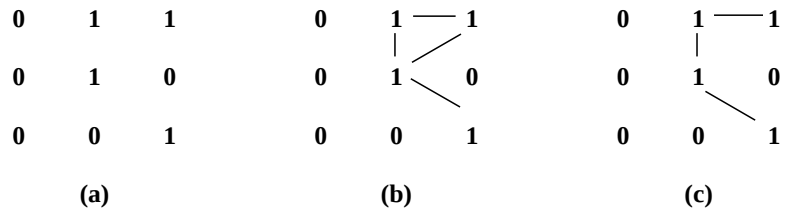
Doi pixeli *p* și *q* cu valori în *V* sunt m-conectați dacă:

(1) *q* este în $N_4(p)$

sau

(2) *q* este în $N_D(p)$ **și** mulțimea $N_4(p) \cap N_4(q)$ este vidă (aceasta este mulțimea pixelilor care sunt vecini în N_4 atât pentru *p*, cât și pentru *q* și ale căror valori sunt în *V*).

Conectivitatea mixtă a fost introdusă ca o modificare a conectivității 8 în scopul eliminării conexiunilor ce conduc la căi multiple și care apar deseori atunci când se utilizează conectivitate 8 - exemplul (b) prezentat în continuare:



(a) **(b)** **(c)**

Definiții

Un pixel *p* este **adiacent** unui pixel *q* dacă ei sunt conectați.
Două subseturi S_1 şi S_2 ale imaginii sunt adiacente dacă există cel puțin un pixel din S_1 adiacent cu un altul din S_2.

Un **drum** de la un pixel *p* cu coord. *(x,y)* la un pixel *q* cu coord. *(s,t)* este o secvență de pixeli distincți cu coordonatele

$(x_0,y_0), (x_1,y_1), ..., (x_n,y_n)$
$(x_0,y_0)= (x,y)$ si $(x_n,y_n)= (s,t)$
(x_i,y_i) este adiacent lui (x_{i-1},y_{i-1}), cu $1 \le i \le n$; n = **lungimea drumului** dintre *p* și *q*.

Dacă *p* şi *q* sunt pixeli dintr-un subset *S* al imaginii, atunci ***p* este conectat cu *q* în *S*** dacă există un drum de la *p* la *q* conținut în întregime în *S*.

Pentru orice pixel *p* din *S*, mulțimea pixelilor din *S* conectați cu *p* este numită **componenta conectată** a lui *S*.

Noțiunile de *conectivitate*, *adiacentă* şi *drum* sunt necesare la stabilirea proprietăților unor obiecte din imagini (de exemplu la segmentare, calcul de distanțe, arii şi perimetre, etc.). Anumite pachete software de prelucrari de imagini oferă flexibilitate în stabilirea unor parametri la procesare.

Măsurarea distanței
Considerăm *p, q* şi *z* cu coord. *(x,y), (s,t)* şi *(u,v)*
D este o **funcție distantă** sau o **metrică** dacă:

(1) $D(p,q) \ge 0$ cu $D(p,q) = 0$ dacă $p=q$
(2) $D(p,q)=D(q,p)$
(3) $D(p,z) \le D(p,q) + D(q,z)$

Distanța euclideană

$$D_e(p,q) = [(x-s)^2 + (y-t)^2]^{1/2}$$

Deoarece *x, y, s, t* sunt de regulă numere întregi, aplicarea distanței euclidiene este nepractică în prelucrarea imaginilor digitale. De aceea s-au definit alte tipuri de distanță, care au ca rezultat tot numere întregi.

Distanța D_4

$$D_4(p,q) = |x-s| + |y-t|$$

Exemplu: configurația de pixeli pentru care distanța față de *(x,y)* este $D_4 \le 2$

		2		
	2	**1**	**2**	
2	**1**	**0**	**1**	**2**
	2	**1**	**2**	
		2		

Pixelii cu $D_4 \le 1$ sunt în $N_4(x,y)$.

Distanța *D8*

$$D_8(p,q) = \max(|x-s|, |y-t|)$$

Exemplu: configurația de pixeli pentru care distanța față de *(x,y)* este $D_8 \leq 2$) *va fi:*

2	**2**	**2**	**2**	**2**
2	**1**	**1**	**1**	**2**
2	**1**	**0**	**1**	**2**
2	**1**	**1**	**1**	**2**
2	**2**	**2**	**2**	**2**

Pixelii cu $D_8 \leq 1$ sunt în $N_8(x,y)$.

Lungimea drumului D_4 între două puncte este egală cu lungimea celui mai scurt drum între cele două puncte cu respectarea convențiilor pentru *drum* şi *distanță*. Acest lucru se aplică şi pentru D_8.

Operații aritmetice şi logice

Operațiile aritmetice şi logice sunt utilizate pe scara largă în prelucrarea imaginilor. Operațiile logice se aplică doar pe imagini binare, în timp ce operațiile aritmetice se aplică pe imagini cu mai multe nivele de gri.

Operațiile aritmetice dintre doi pixeli p şi q sunt notate astfel:

adunare: $p+q$
scădere: $p-q$
înmulțire: $p*q$ (sau pq sau $p \times q$)
împărțire: $p \div q$

Operațiile aritmetice pe imagini întregi sunt efectuate pixel cu pixel. Principala utilizare a adunării este aceea de mediere a imaginii în scopul reducerii zgomotului. Scaderea imaginilor este utilizată mult în imagistica medicală, ca instrument prin care se înlatură informațiile date de fondul static al imaginilor prelucrate (atunci când acesta se cunoaşte). Înmulțirea imaginilor (sau împărțirea) se utilizează pentru corecția umbrelor ce provin din neuniformități în iluminare sau în sensibilitatea senzorului utilizat pentru achiziția imaginii.

Operațiile aritmetice implică locația spațială a unui singur pixel la un moment dat, de aceea pot fi realizate "pe loc": rezultatul efectuării operației la locația *(x,y)* poate fi memorată la acea locație într-una din imaginile existente (participante la operația aritmetică), deoarece locația respectivă nu va mai fi "vizitată" a doua oară.

Operațiile logice utilizate în procesarea imaginilor sunt AND (“şi” logic), OR (“sau” logic) şi COMPLEMENT (negare logică), notate astfel:

AND:	*p AND q* (sau $p \cdot q$)
OR:	*p OR q* (sau $p+q$)
COMPLEMENT:	*NOT p* (sau $\sim p$)

Aceste operații sunt complete din punct de vedere funcțional (pot fi combinate pentru a obține orice altă operație logică).

Operațiile logice sunt instrumente de bază în prelucrarea imaginilor deoarece ele sunt utilizate pentru sarcini ca: mascarea, detecția caracteristicilor ("*feature detection*")

şi analiza formelor. Ele sunt executate pixel cu pixel şi, la fel ca în cazul operaţiilor aritmetice, pot fi executate "pe loc".

Figura III.2.6.3 prezintă câteva exemple de operaţii logice pe imagini binare.

Pe lângă prelucrarea unor întregi imagini, pixel cu pixel, operaţiile aritmetice şi logice sunt utilizate şi în operaţii orientate pe o vecinatate. Procesarea pe o vecinatate este formulată în contextul aşa-numitelor **operaţii cu "mască"** - termenii de "tipar" ("*template*"), "fereastră" sau "filtru" sunt folosiţi pentru aceeaşi noţiune.

Ideea care stă la baza acetui tip de operaţii este de a permite ca valoarea unui pixel sa fie stabilită ca o funcţie dependentă de nivelul de gri iniţial pentru pixelul respectiv, împreună cu nivelele vecinilor acestuia.

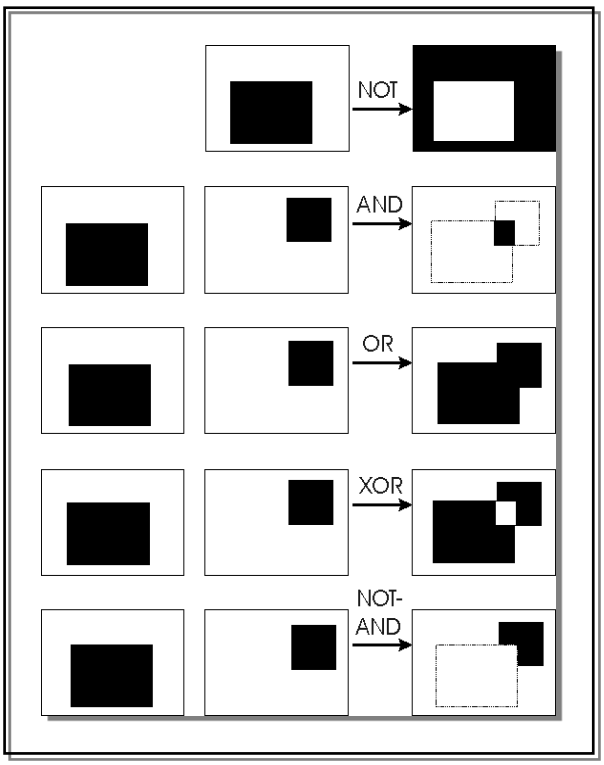


Figura III.2.6.3. Exemple de operaţii logice pe imagini binare. De notat faptul că negrul reprezintă în cazul acesta valoarea logică “adevarat” (sau 1)

Ca exemplu, să considerăm o subimagine (a) în care dorim să înlocuim valoarea lui z_5 cu media pixelilor dintr-o regiune de 3*3 pixeli centraţi în z_5, adică:

$$z = \frac{1}{9}(z_1 + z_2 + ... + z_9) = \frac{1}{9}\sum_{i=1}^{9} z_i$$

z1	***z2***	***z3***	***w1***	***w2***	***w3***
z4	***z5***	***z6***	***w4***	***w5***	***w6***
z7	***z8***	***z9***	***w7***	***w8***	***w9***
	(a)			(b)	

Dacă introducem masca de ponderare (b), vom putea nuanța "importanța" acordată valorilor de gri inițiale pentru cei nouă pixeli din vecinătatea considerată: o nouă valoare a lui z_5 va fi *media ponderată*:

$$z = w_1 z_1 + w_2 z_2 + ... + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

Situația precedentă (z_5 media aritmetică) se va regăsi ca un caz particular în care ponderile sunt toate egale: w_i=1/9, i=1,2,...,9.

Geometria imaginilor

Vom prezenta în cele ce urmează doar transformări de bază, fără a intra în probleme legate de transformările de perspectivă.

Materialul acestei sectiuni urmareşte doar să dea o idee generală referitoare la formalizarea unor probleme ca translația, scalarea sau rotirea imaginilor.

Toate transformările sunt prezentate în sistemul de coordonate cartezian tri-dimensional (3-D), în care coordonatele sunt notate cu *(X,Y,Z)*. În mod uzual, coordonatele *(X,Y,Z)* se numesc "*world coordinates*".

Translația

Să presupunem că ne propunem să translatăm un punct de coordonate *(X,Y,Z)* într-o nouă locație utilizând deplasamentele (X_0,Y_0,Z_0). Translația va fi realizată utilizând ecuațiile următoare (cea din dreapta reprezintă forma matricială):

$$\begin{cases} X^* = X + X_0 \\ Y^* = Y + Y_0 \\ Z^* = Z + Z_0 \end{cases} \qquad \begin{bmatrix} X^* \\ Y^* \\ Z^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

unde (X^*,Y^*,Z^*) sunt coordonatele noului punct, iar T se numeşte *matrice de transformare*.

Se preferă utilizarea matricilor pătrate:

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad \text{iar} \quad v^* = Tv$$

În mod similar se procedează pentru alte transformări geometrice (scalarea, rotirea). Este deseori util să se utilizeze mai multe transformari pentru a se produce un rezultat compus, de exemplu translație, urmată de scalare şi apoi de rotație. Utilizarea matricilor pătrate simplifică mult reprezentarea formală a acestui proces deoarece concatenarea transformărilor se face prin compunerea matricilor şi obținerea unei **matrici de transformare unice**.

Pentru transformările inverse se determină inversele matricilor de transformare (inversele matricilor corespunzatoare transformărilor mai complexe se obțin prin tehnici numerice).

2.7. ÎMBUNĂTĂȚIREA IMAGINILOR ŞI EXTRAGEREA UNOR ATRIBUTE

Principalul obiectiv al tehnicilor de îmbunătățire este acela de a prelucra o imagine în aşa fel încat rezultatul sa fie mai potrivit decât imaginea inițială pentru o anumită aplicație – spunem că sunt *orientate pe problemă* ("*problem-oriented*").

De exemplu, o metoda foarte potrivită pentru îmbunătățirea imaginilor obținute cu raze X poate să nu fie cea mai potrivită abordare în cazul imaginilor ecografice sau în prelucrarea imaginilor transmise de pe Marte.

Abordările discutate în acest capitol intra în două categorii largi (analoage metodelor de la prelucrările de semnale): metode în domeniul spațial şi metode în domeniul frecvențial.

Domeniul spațial se referă la planul imaginii, iar metodele din această categorie sunt bazate pe manipularea directă a pixelilor din imagine.

Domeniul fecvențial cuprinde tehnicile bazate pe modificarea transformatei Fourier a imaginii.

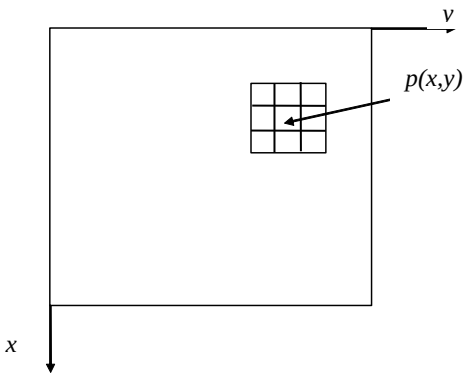
În cele ce urmează vom aborda doar metode din domeniul spațial.

Metode spațiale

Funcțiile de acest tip pot fi exprimate ca fiind:

$$g(x,y) = T\left[f(x,y)\right]$$

unde $f(x,y)$ este imaginea initială
$g(x,y)$ imaginea procesată
T este un operator pe f, definit pe o vecinatate a lui (x,y)



Centrul subimaginii este mutat pixel cu pixel pornind (de exemplu) din colțul stânga sus şi aplicând operatorul T pentru fiecare locație. Aplicațiile de prelucrări de imagini utilizează tehnici de optimizare care reduc considerabil timpul de prelucrare.


(a)

(b)

Figura III.2.7.1. Negativul unei imagini – o transformare punctuală (noua valuare este în funcție doar de valoarea de gri inițială și nu este influențată de vecini). Prelucrare cu *Image-Pro Plus v.3.0*

Îmbunătățirea prin procesare punctuală

Transformări asupra intensității – au ca formulă generală:

$$s=T(r)$$

unde *r* este valoarea de gri inițială, iar *s* este rezultatul aplicării operatorului *T*.
La procesările punctuale, noua valoare este funcție doar de valoarea de gri inițială şi nu este influențată de vecini.

- **negativul unei imagin**i – figura III.2.7.1.
- **întărirea contrastului** ("*stretching*") – figura III.2.7.2.

Figura III.2.7.2. Întărirea contrastului ca transformare punctuală (aceeași imagine inițială ca cea din figura 7.2 a). Imagine prelucrată cu *Image-Pro Plus v.3.0*

- **Împărțirea în plane binare**

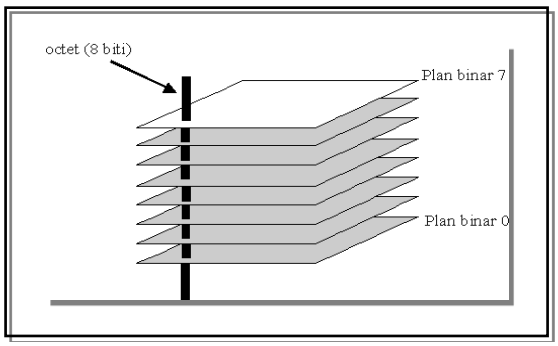


Figura III.2.7.3. Împărțirea unei imagini cu 256 nivele de gri în 8 plane binare

Să presupunem că valoarea fiecărui pixel din imagine este reprezentată pe 8 biți (1 octet). Să ne imaginăm că imaginea este compusă din 8 plane binare (cu valori reprezentate pe 1 bit) astfel încât fiecare plan conține biții de un anumit rang - de la planul 0 (cel mai puțin semnificativ) la planul 7 (cel mai semnificativ). Figura III.2.7.3 ilustrează această împărțire a unei imagini în plane binare numite "*bitplanes*".

Figura III.2.7.4 prezintă o imagine şi planele ei binare, începând cu cel mai puțin semnificativ. Observați că doar datele din planele binare de ordin înalt (cele 4-5 plane corespunzătoare biților cei mai semnificativi - planele [3,] 4, 5, 6, 7) conțin informație perceptibilă, celelalte plane conțin detalii de subtilitate.

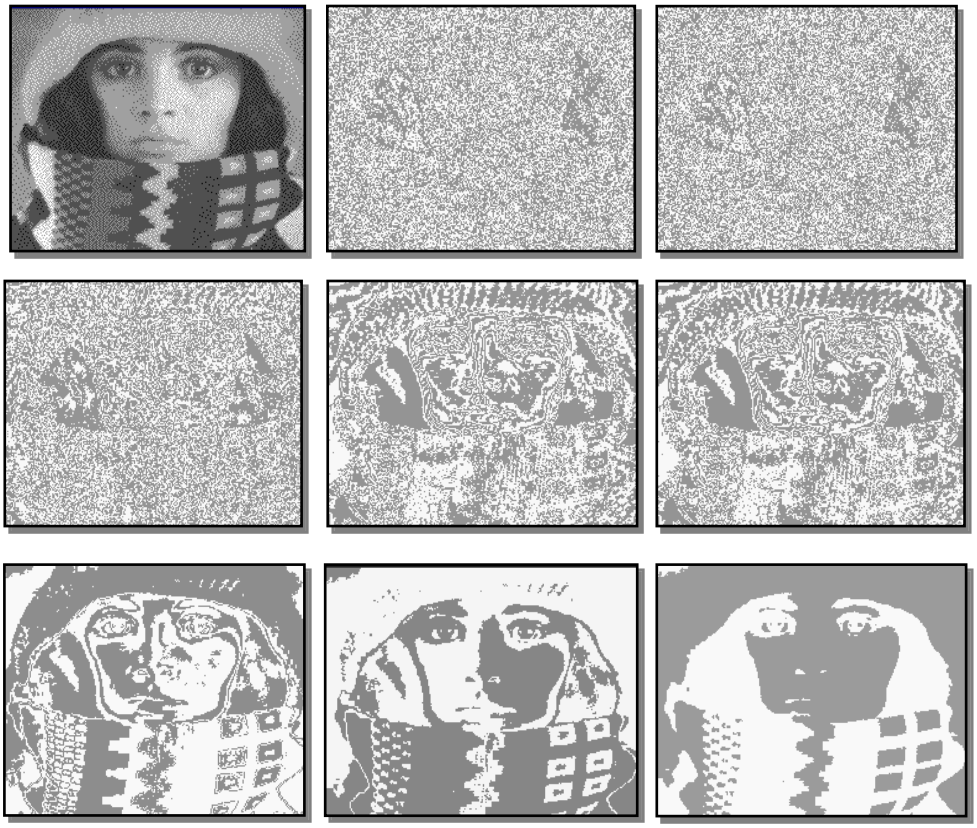


Figura III.2.7.4. Planele binare ale unei imagini şi modul cum este "distribuită" informația utilă din imagine între planele de rând diferit. Imagine preluată şi prelucrată cu *AIM-Another Image Manager* (Universitatea din Amsterdam)

Procesarea histogramei

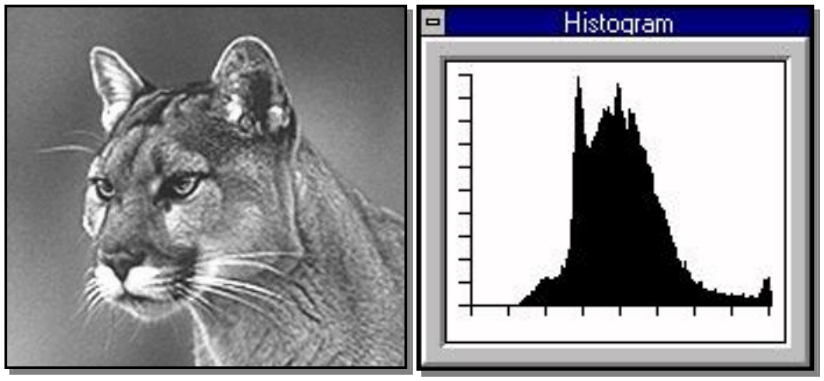
Histograma unei imagini digitale cu L nivele de gri (în domeniul [0 , L-1]) este o funcție discretă:

$$p(r_k) = \frac{n_k}{n}$$

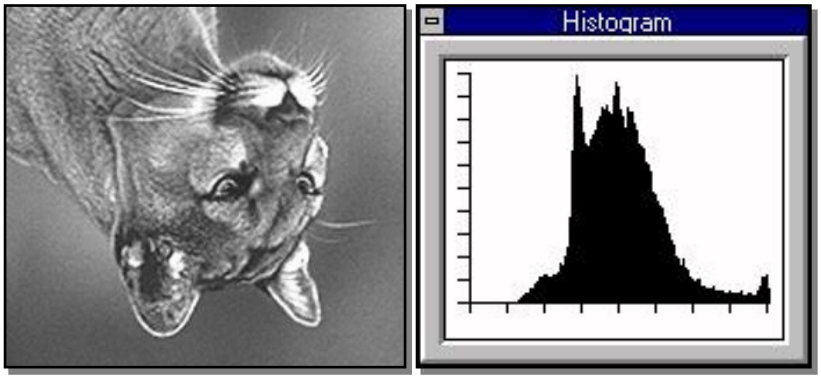
r_k - nivelul de gri de ordinul k, cu k=0, 1,2, ..., L-1
n_k - numărul de pixeli cu nivelul de gri de ordinul k
n - numărul total de pixeli din imagine

Putem spune că **histograma $p(r_k)$** ne dă o estimare a probabilității de apariție a nivelului de gri r_k în imaginea studiată.

O reprezentare grafică a acestei funcții ne dă o descriere globală a aspectului imaginii. Deşi histograma unei imagini este o descriere care nu furnizează informație privind conținutul concret al unei imagini (figura III.2.7.5 ilustrează acest lucru), alura histogramei unei imagini aduce informație deosebit de prețioasă privind posibilitatea îmbunătățirii contrastului.



(a)

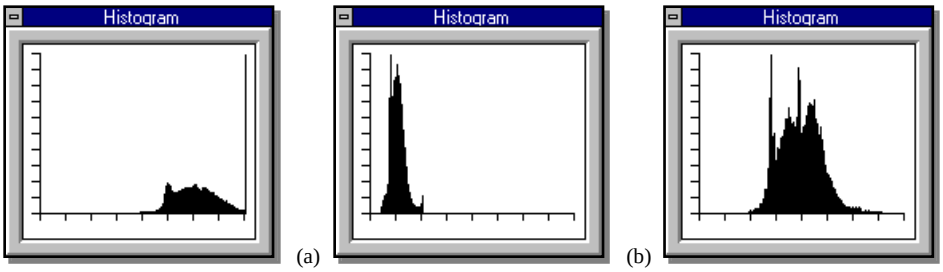


(b)

Figura III.2.7.5. Histograma ne dă informații privind utilizarea nivelelor de gri, nu conținutul imaginii (imaginea şi răsturnata ei au aceeaşi histogramă)

Figura III.2.7.6 prezintă histogramele pentru trei tipuri de imagini: o imagine (a) cu un contrast nu foarte bun (observați aglomerarea nivelelor de gri în centrul intervalului de valori, lucru deseori acceptabil pentru aplicații uzuale) şi două histograme ale unor imagini de slabă calitate - una foarte întunecată (b) şi una extrem de luminoasă (c).

O metodă de îmbunătățire bazată pe utilizarea histogramei şi larg utilizată pentru anumite tipuri de imagini este cea de **egalizare a histogramei** (ilustrată în figura III.2.7.7), în fapt o egalizare/redistribuire a utilizării nivelelor de gri în imagine.

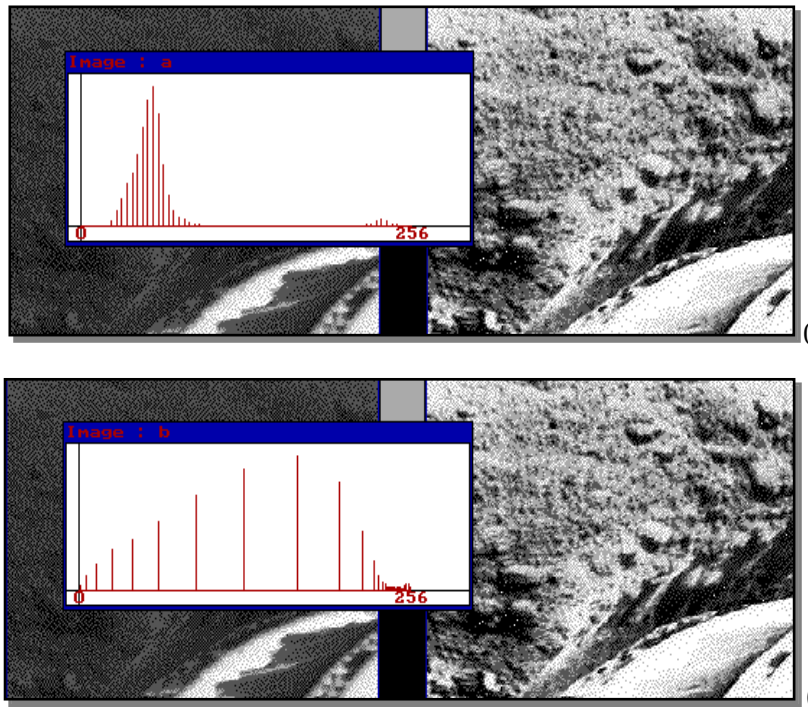


(a) (b) (c)

Figura III.2.7.6. Histograme pentru imagini cu un contrast scăzut (o utilizare ne-eficientă a plajei de gri): imagine ce utilizează doar nivelele de la mijlocul plajei (a); imagine foarte întunecată (b); imagine foarte luminoasă/spălăcită (c)

Principiul metodei urmăreşte o uniformizare a densităţii de probabilitate pe domeniul nivelelor de gri - o repartizare uniformă a acestora în domeniul [0,L-1].

Figura III.2.7.7 prezintă rezultate obţinute aplicând această metodă la prelucrarea unor imagini de pe Marte.



(a)

(b)

Figura III.2.7.7. Efectul egalizării histogramei – informaţiile ies la iveală. Histograma (a) pentru imaginea iniţială din stânga arată utilizarea aproape exclusivă a nuanţelor foarte închise şi aşa de apropiate pe scara de gri încât sistemul vizual uman nu le poate distinge. Nivelele de gri sunt redistribuite pe plaja disponibilă (b), astfel încât distanţa dintre două valori consecutive creşte şi ele pot fi percepute ca nuanţe distincte (imaginea din dreapta). Imagine preluată şi prelucrată cu *AIM-Another Image Manager* (Universitatea din Amsterdam)

Filtrarea spațială

Filtrele liniare folosesc o "mască" (b) pentru determinarea unei noi valori a pixelului z_5 din centrul unei vecinătăți (a) ca o relație liniară între valorile de gri inițiale z_i şi ponderile w_i ce exprimă relația dorită între pixeli.:

$$R = w_1 z_1 + w_2 z_2 + ... + w_9 z_9 = \sum_{i=1}^{9} w_i z_i$$

z1	***z2***	***z3***	***w1***	***w2***	***w3***
z4	***z5***	***z6***	***w4***	***w5***	***w6***
z7	***z8***	***z9***	***w7***	***w8***	***w9***
	(a)			(b)	

Filtrele neliniare utilizează funcții neliniare pentru determinarea noii valori din centrul unei vecinătăți – de exemplu, pentru valoarea maximă din centrul unei vecinătăți de 3*3 pixeli:

$$R = \max \{ z_k \mid k = 1,2,....,9 \}$$

Similar se utilizează şi alte funcții: valoarea minimă, mediana, etc.

Filtre pentru netezire

Acest tip de filtrare este prezentat în figura III.2.7.8 – are efect de încețoșare şi voalare a contururilor (*blurring*). De aceea, filtrele de netezire se mai numesc ”filtre integrative”. În cazul filtrelor liniare, ponderile sunt pozitive la acest tip de filtrare.

Observați că cele trei metode de netezire din figura III.2.7.8 nu dau rezultate identice pe cele două imagini: la imaginea cu zgomot uniform, filtrarea mediană nu pare să dea rezultate spectaculoase, pe cand la imaginea afectată de zgomot binar filtrarea mediană pare cea mai performantă. Explicația se găseşte în modul în care este ales indicatorul tendinței centrale pentru cele două tipuri de distribuții ale zgomotului (vezi capitolul de prelucrări statistice).

Aici s-au ilustrat doar situații simple, în care doar distribuția valorilor zgomotului diferă (distribuție uniformă, respectiv binară), utilizându-se o distribuție spațială uniformă. Prelucrările devin mai complexe când distribuția pixelilor zgomotoşi este neuniformă spațial; în acelaşi timp, valorilor de gri pot urma diverse distribuții, mai complexe decât cele ilustrate aici.

Filtre derivative

Filtrele derivative utilizează măşti cu ponderi negative şi pozitive care conduc la o “derivare” a imaginii, punând în evidență tranzițiile şi schimbările de comportament ale funcției imagine (detaliile de finețe şi tranzițiile între nivelele de gri).

Figura III.2.7.9 prezintă rezultatul aplicării unui filtru de tip gradient (o derivată de ordinul I), iar figura III.2.7.10 a unui filtru Laplace (o derivată de ordinul II).

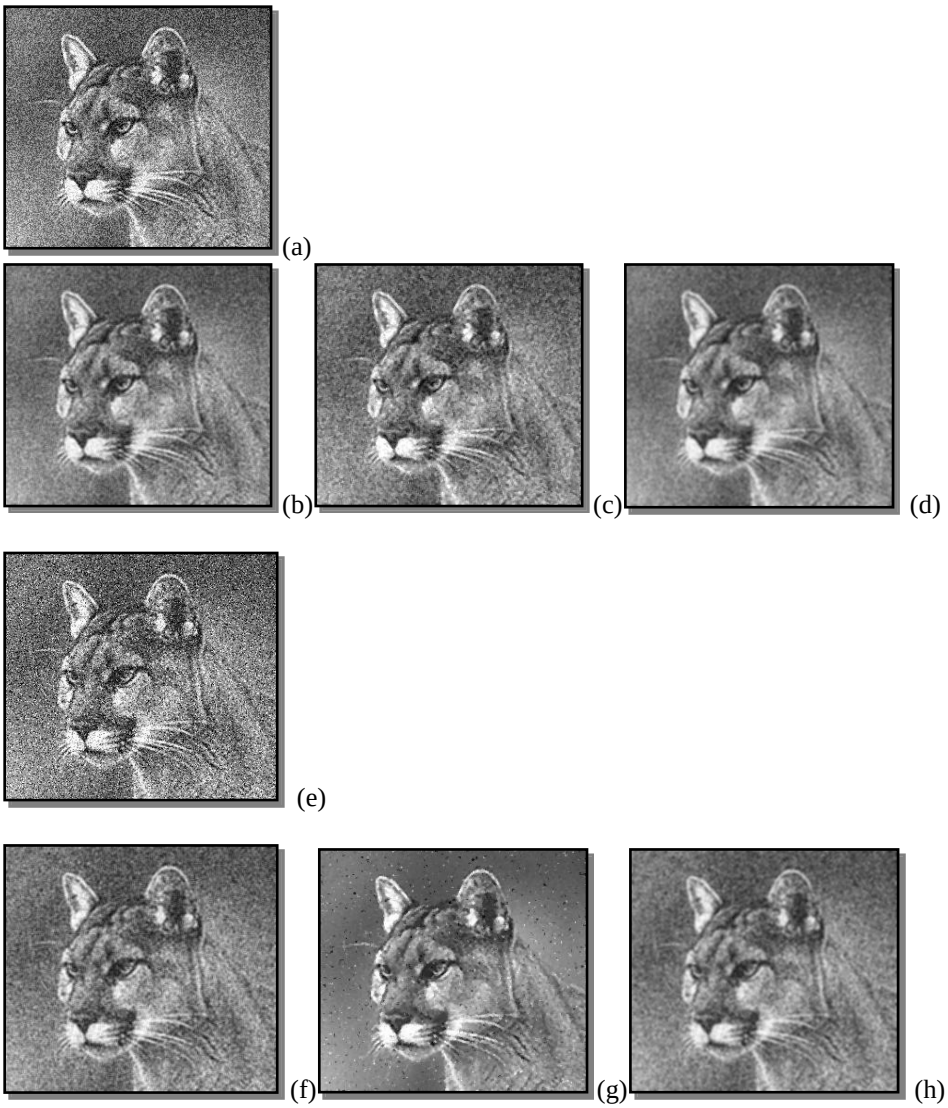
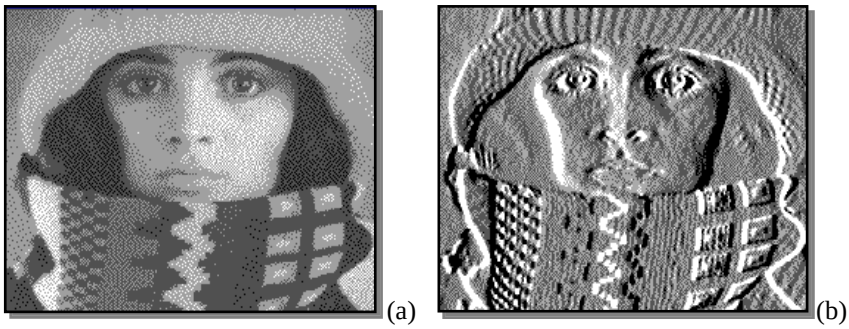
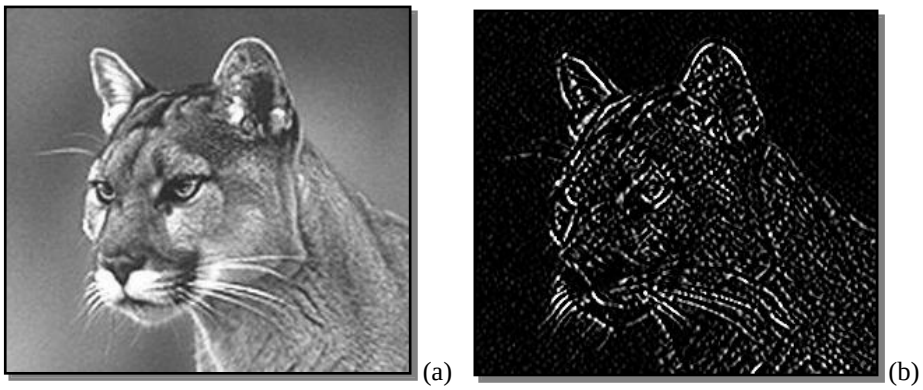


Figura III.2.7.8. Efectul filtrelor de netezire. Imaginile (a) şi (e) sunt cele originale, cu acelaşi nivel de zgomot adăugat în mod uniform din punct de vedere spațial – imaginii (a) i s-a adăugat zgomot cu valori între 0 şi 255 distribuit în mod uniform pe plaja de gri (adică toate nuanțele de gri), în timp ce imaginii (e) i s-a suprapus zgomot binar, de tip "sare şi piper" (doar extremele de 0 şi 255). Filtrele aplicate au fost: netezire uniformă în (b) şi (f); netezire Gausiana în (d) şi (h); filtrare mediană în (c) şi (g). Imagine prelucrată cu *Image-Pro Plus v.3.0*



(a) (b)

Figura III.2.7.9. Rezultatul aplicării unui filtru gradient. Imagine preluată şi prelucrată cu *AIM-Another Image Manager* (Universitatea din Amsterdam)



(a) (b)

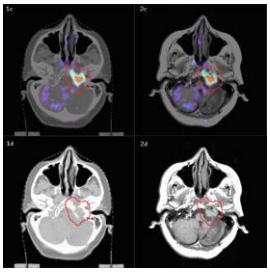
Figura III.2.7.10. Rezultatul aplicării unui filtru Laplace. Imagine prelucrata cu *Image-Pro Plus v.3.0*

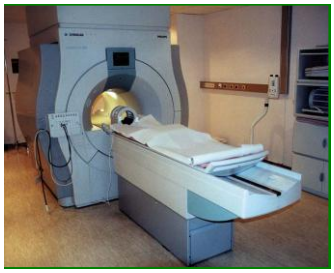
Se observă modul diferit de evidențiere a muchiilor pentru cele două tipuri de filtre derivative (de ordinul I sau II).

Prezentarea unor probleme elementare din prelucrarea imaginilor a avut scopul de a ajuta la formarea unei idei generale privind problematica domeniului. Totodată, credem că aceste noțiuni de bază vor constitui fundamentul necesar atât pentru a putea alege unele din opțiunile oferite de programele de prelucrare, cat şi pentru studierea mai aprofundată a problemelor specifice imagisticii medicale.

2.8. STANDARDUL DICOM

Digital Imaging and Communications in Medicine





Standardul **DICOM** facilitează interoperabilitatea echipamentelor de imagistică medicală — standardul specifică:

- un set de protocoale ce trebuiesc respectate de către toate echipamentele care pretind ca sunt conforme standardului
- sintaxa şi semantica comenzilor, precum şi informaţia asociată acestor protocoale
- informaţiile ce trebuiesc furnizate de către echipamentele ce se conformează standardului

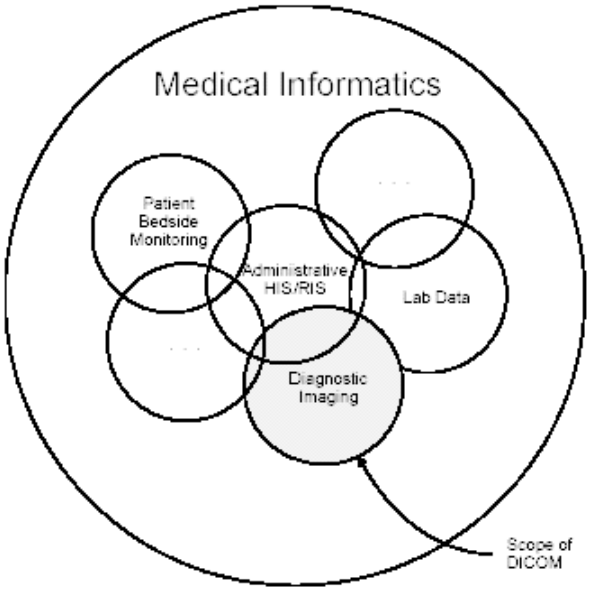


Figura III.2.8.1. Datele furnizate de echipamentele de imagistică medicală trebuie să se integreze în circuitul general al informaţiei medicale. [DICOM]

Scurt istoric

- anii 1970 — introducerea tomografiei computerizate, urmată de dezvoltarea altor tehnici de investigare imagistică — nevoia unor standarde de transfer a imaginilor şi informaţiei asociate acestora între echipamentele furnizate de diverşi producători
- 1983 — *American College of Radiology* (**ACR**) şi *National Electrical Manufacturers Association* (**NEMA**) formează un comitet care dezvoltă standardul **DICOM** (dezvoltat şi publicat potrivit standardelor **NEMA** şi în acord cu directivele ISO/IEC)
  Standardul a fost dezvoltat împreună cu alte organizaţii internaţionale de standardizare
  CEN TC251 – Europa
  JIRA Japonia
  IEEE
  HL7
  ANSI - SUA
- 1988 – **DICOM** versiunea 2
- 2001 – **DICOM** versiunea 3 (publicată de **NEMA**).

Standardul **DICOM v.3** este aplicabil în reţele (respectă protocoalele standard de reţea OSI şi TCP/IP) realizând interoperabilitatea completă între sistemele conectate în reţea (figura III.2.8.2).

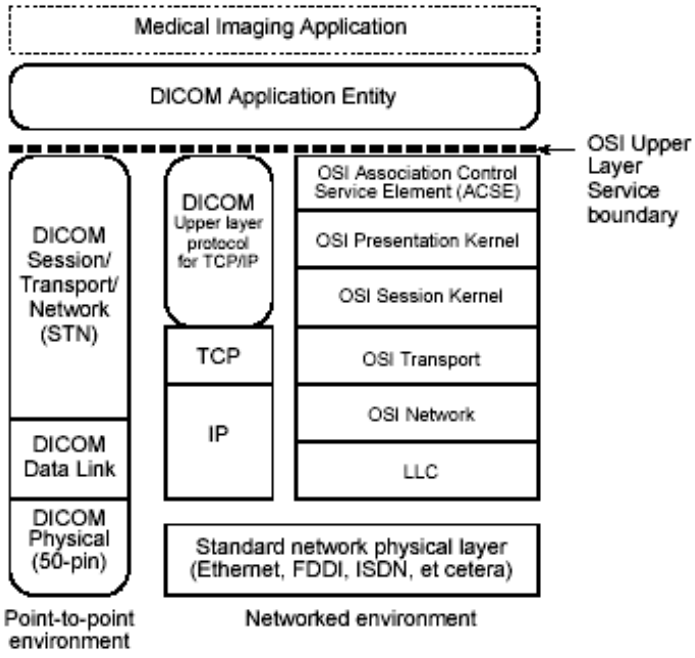


Figura III.2.8.2. Standardul DICOM permite conectarea echipamentelor de imagistică medicală în reţea şi interoperabilitatea cu celelalte echipamente medicale. [DICOM]

DICOM versiunea 3:

- specifică clar clasele de servicii, semantica comenzilor şi tipurile de date ⇒ modul în care trebuie să reacționeze dispozitivele care pretind că respectă standardul
- specifică nivele de conformare cu standardul
- are organizare modulară – se pot adăuga noi facilități
- introduce în mod explicit "*Information Objects*" nu numai pentru imagini şi grafică, ci şi pentru studii, rapoarte, etc.
- precizează tehnica pentru identificarea neambiguă a relațiilor dintre diferitele informatii ("*Information Objects*") din rețea.

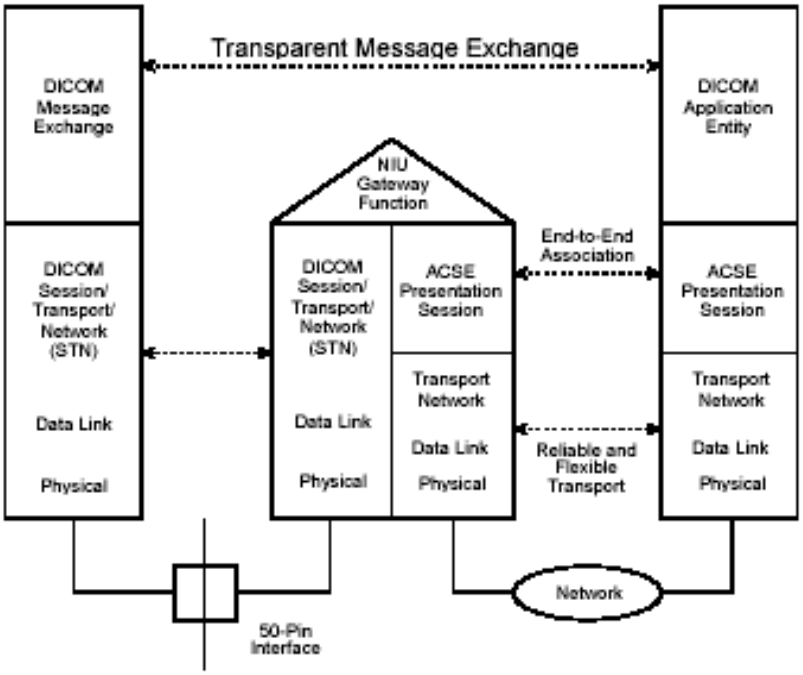


Figura III.2.8.3. Standardul DICOM are organizare modulară şi oferă flexibilitate, conectând în mod natural imaginile cu celelate informații medicale. [DICOM]

Mai multe informații se pot afla de pe diferite site-uri specializate – o bună sursă este *Penn State Radiology*.

BIBLIOGRAFIE ŞI REFERINȚE

I. Bankman. *Handbook of medical imaging. Processing and analysis management*. Academic Press, San Diego, 2000

JH van Bemmel, MA Musen (eds). *Handbook of Medical Informatics*. Springer Verlag, Heidelberg, 1997

DICOM. Digital Imaging and communications in Medicine: *http://medical.nema.org/*

RC Gonzalez, RE Woods. *Digital Image Processing* (2nd ed.). Prentice Hall, Englewood Cliffs, NJ, 2002

National Library of Medicine. The Visible Human Project: *www.nlm.nih.gov/research/visible/visible_human.html*

Penn State Radiology: *http://www.xray.hmc.psu.edu/physresources/dicom*/

O. Popescu (ed). *Informatica medicala*. Editura Medicala, Bucureşti, 1988

Partea a IV-a

DECIZIA MEDICALĂ ASISTATĂ DE CALCULATOR

INTRODUCERE

Asistăm azi la extinderea utilizării calculatoarelor în cele mai diverse domenii, medicina fiind un domeniu vizat încă de la apariția primelor calculatoare. Atracția pentru medicină nu este întâmplătoare: informaticienii primelor generații de calculatoare simțeau nevoia de a demonstra universalitatea aplicațiilor şi doreau exemple în afara granițelor tradiționale ale disciplinelor "exacte" în care formalizarea matematică lasă să se întrevadă uşor aplicațiile. Pentru un nespecialist, activitatea medicului la stabilirea diagnosticului pare un simplu proces de asociere a unui termen, numit "diagnostic", cu un set de elemente numite "simptome". În primele programe de diagnostic asistat s-a procedat similar, însă rezultatele obținute au fost deosebit de modeste. La o analiză mai atentă se poate vedea că nici nu se puteau aştepta rezultate mai bune într-o viziune atât de simplistă. Multe semne nu pot fi încadrate într-o logică bivalentă, nu toate semnele apar cu necesitate, nu toate au aceeaşi valoare în stabilirea diagnosticului; în plus, există frecvente complicații sau asocieri care fac ca mulțimea diagnosticelor să fie greu de definit. Peste toate acestea s-a adăugat şi constatarea că raționamentul medicului se deosebeşte substanțial de raționamentul liniar deductiv-exclusiv, fiind considerat oarecum mai apropiat de procesul de recunoaştere. Actualmente raționamentul medical se găseşte sub lupa specialiştilor în inteligența artificială şi reprezentarea cunoştințelor, ca şi a celor interesați în ştiințe cognitive. S-a progresat foarte mult în această direcție, cadrul de dezvoltare, mult mai larg decât cel strict limitat la "diagnosticul asistat", fiind oferit de logica matematică, ce permite formalizarea cunoştințelor în general. Apariția limbajelor logice - de exemplu PROLOG - a deschis efectiv un nou mare capitol al informaticii medicale: "decizia medicală asistată de calculator", care alături de aplicațiile privind prelucrarea semnalelor şi imaginilor, bazele de date şi sistemele informatice medicale, biostatistică şi modelarea proceselor biologice, constituie domenii care îşi schimbă încetul cu încetul statutul din domenii de vârf în domenii de rutină. Ritmul ridicat în care apar noi programe în aceste discipline face însă ca să fie tot mai dificil de ținut pasul cu noutățile mai ales când ele cu greu îşi fac loc în programele tradiționale ale învățământului medical.

Aplicațiile practice privind asistarea deciziei medicale nu s-au limitat la cele privind diagnosticul asistat, acoperind şi alte aspecte. Putem distinge astefel următoarele direcții:

a) diagnostic asistat
b) alegerea investigațiilor
c) optimizarea tratamentului
d) asistarea deciziei în managementul sanitar.

Vom analiza pe rând aceste aspecte, tratând mai detaliat aplicațiile în domeniul diagnosticului asistat.

1. DIAGNOSTICUL ASISTAT DE CALCULATOR

1.1. CLASIFICAREA METODELOR DE DIAGNOSTIC

Rezultatele modeste ale primelor programe de diagnostic "automat" (denumire la care actualmente s-a renunţat) nu au determinat însă o a abandonare a temei, ci tocmai o creştere a eforturilor de analiză aprofundată a raţionamentului medical, cu ecou până la dezvoltarea unor abordări teoretice de finețe. Se pot distinge trei direcţii principale de studiu:

- îmbunătăţirea ***metodelor logice*** prin care s-a pornit iniţial abordarea diagnosticului asistat
- elaborarea unor ***metode statistice*** adecvate, conducând până la teoria clasificării
- formalizarea cunoştinţelor medicale în forma lor ***euristică***, cu scopul direct declarat de simulare a raţionamentului medical, conducând la elaborarea sistemelor expert.

În continuare vom prezenta stadiul actual al cercetărilor în acest domeniu şi vom sistematiza metodele abordate, deşi unele metode au caracter hibrid şi sunt greu de clasificat. Totodată, prezentarea se va face evitând abordarea teoretică, încercând a expune într-o manieră descriptivă cele mai importante direcţii de dezvoltare a "diagnosticului asistat de calculator".

1.2. FORMALIZAREA OPERAŢIUNII DE STABILIRE A DIAGNOSTICULUI

Operaţiunea de stabilire a diagnosticului poate fi privită ca rezultatul confruntării comparative de către medic a două fluxuri de informaţie:

- informaţii privind starea pacientului, obţinute atât prin dialog direct (anamneză) cât şi prin diverse investigaţii (laborator, radiografii, explorări funcţionale etc.): ansamblul acestor informaţii, cu valori concrete pentru pacientul investigat (simptome) le numim ***date.***
- informaţii pe care medicul le posedă (din pregătire, din experienţa clinică, din materiale documentare etc.) şi pe care le numim ***cunoştinţe***, cuprind o multitudine de relaţii între simptome şi diagnostice.

Operaţiunea de stabilire a diagnosticului este prezentată schematic în fig. IV.1.

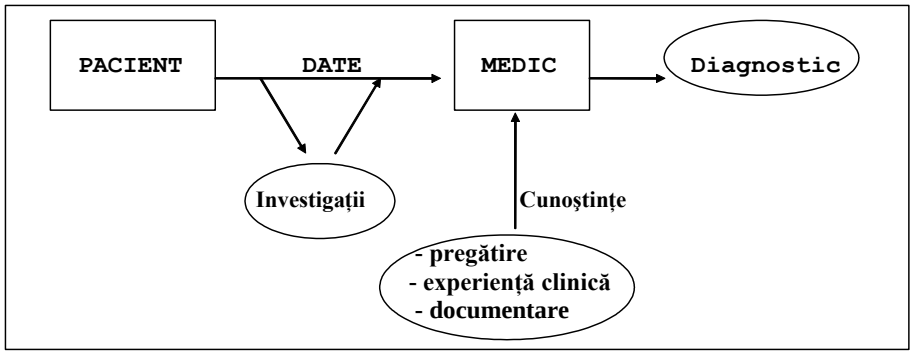


Figura IV.1. Reprezentarea schematică a fluxurilor informaţionale în operaţia de stabilire a diagnosticului

Realizarea unor programe de calculator pentru "diagnostic asistat" trebuie să urmărească deci confruntarea celor două fluxuri de informaţii:

- cunoştinţele din domeniu, care formează un ansamblu numit ***baza de cunoştinţe (BC)***

- datele despre pacient, anasamblu numit şi ***vectorul de stare*** al pacientului (***PAC***).

Metodele de diagnostic enumerate mai sus se deosebesc între ele prin modul în care este construită baza de cunoştinţe, modul de culegere (reprezentare) a datelor pacientului şi modul de confruntare date-cunoştinţe (raţionament).

Să le abordăm pe rând.

2. METODE LOGICE

2.1. BAZA DE CUNOŞTINE

Metodele logice folosesc drept bază de cunoştinţe o matrice boli/simptome (Tabel IV.1).

Tabel IV.1. Structura bazei de cunoştinţe pentru metodele logice în diagnosticul asistat

Simptome / Diagnostice	S_1	S_2	S_3	...	S_m
D_1 (Hipertiroidism)	0	1	0(n)		**1**
D_2 (Hipertensiune)	1	0	1		0(n)
. . .	. . .	. . .	. . .	. . .	. . .
D_n (Angină pectorală)	1	1	0		0(n)

D_1, D_2 ... D_n reprezintă diagnostice iar S_1, S_2 ... S_m reprezintă simptome. Ex: S_1 = hipertrofie ventriculară stângă, S_2 = palpitaţii, S_3 = cefalee,... S_m = exoftalmie

Matricea este construită pe baza logicii bivalente (de unde şi denumirea de metode "logice"), valoarea "1" reprezentând prezenţa simptomului în diagnosticul respectiv, iar valoarea "0" absenţa. (În variantele mai noi se utilizează şi "n" cu semnificaţia "nu are importanţă", adică prezenţa/absenţa simptomului respectiv este nerelevantă pentru acel diagnostic).

Conţinutul acestul tabel este realizat prin contribuţia unui grup de experţi şi este reprezentat în program sub forma unei matrici, fiecare element din matrice fiind caracterizat prin doi indici: unul referitor la boală, celălalt la simptom.

2.2. VARIANTE DE METODE LOGICE

În metodele logice vectorul de stare al pacientului este format dintr-un şir de valori corespunzătoare tuturor simptomelor existente în baza de cunoştinţe. La lansarea programului vectorul PAC conţine "0" în toate căsuţele.

După modul de construcţie al vectorului de stare al pacientului distingem două variante mai însemnate de metode logice.

a) Tabele de adevăr

Simptomele sunt clasificate şi prezentate pe mai multe pagini de ecran. De obicei selecţia se efectuează prin deplasarea pe ecran a unui cursor până la simptomul ales şi apăsarea unei taste de selecţie (de obicei "Enter"); se deplasează apoi cursorul pe altă poziţie şi aşa mai departe. Pentru simptomele selectate vectorul de stare al pacientului va conţine "1", restul căsuţelor rămânând "0". Majoritatea programelor menţin cel puţin o linie "menu" în tot timpul rulării pentru precizarea modului de lucru în secvenţa corespunzătoare din program.

b) Arbori de decizie

Pentru o apropiere mai mare de clasica anamneză, în loc de selectarea de pe ecran a simptomelor, programul afişează succesiv întrebări privind prezenţa / absenţa unor simptome, solicitând răspuns de tip da/nu. În cazul unui răspuns pozitiv, în vectorul de stare al pacientului apare "1" în poziţia simptomului respectiv.

Varianta "arbori de decizie" impune alegerea succesiunii întrebărilor în funcţie de răspunsurile primite, evitarea întrebărilor inutile. (de ex. întrebări despre sarcină la un bărbat etc.). Se apreciază că prin utilizarea arborilor de decizie, lăsând pacientul însuşi să răspundă la unele întrebări se poate obţine o mai puternică implicare a sa, cu rezultate pozitive în tratament.

2.3. PREZENTAREA REZULTATELOR

Indiferent de varianta prin care se construieşte vectorul de stare al pacientului, în continuare acesta este comparat cu fiecare linie a matricii boli/simptome şi se calculează câte din simptomele caracteristice bolii respective sunt prezente la pacient. De exemplu, dacă boala D_1 are 8 simptome caracteristice (8 valori notate cu "1" în BC), iar pacientul nostru prezintă 6 dintre acestea, atunci procentul de simptome ale bolii D_1 prezente la pacient este 6/8 = 75%.

Prezentarea rezultatului poate fi efectuată în mai multe feluri:

a) prin enumerarea tuturor diagnosticelor în care apar semnele pacientului, în lista afiliată diagnosticele sunt ordonate după numărul de simptome care coincid;

b) prin calcularea unui procent de coincidenţe şi ordonarea diagnosticelor după acest procent;

c) prezentarea pentru fiecare diagnostic din lista-rezultat nu numai a numărului de simptome coincidente, ci şi a simptomelor care ar mai fi aşteptate pentru un tablou clinic clasic complet;

d) varianta anterioară îmbunătăţită ar include elemente de diagnostic diferenţial sugerate;

e) prezentarea alături de fiecare diagnostic din lista-rezultat a unor indicaţii terapeutice.

2.4. DEZAVANTAJELE METODELOR LOGICE

Metodele logice, datorită simplităţii lor prezintă o atracţie pentru utilizatori, mai ales metoda tabelelor de adevăr care permite şi o rulare operativă. Totuşi, simplificările, uneori excesive, restrâng utilitatea practică împingând aplicaţiile preponderent în sfera învăţământului medical. Obiecţiile ce pot fi aduse se referă mai mult la două aspecte:

- logica bivalentă de genul "simptomul este prezent/absent" exclude factorul intensitate; într-adevăr, o serie de simptome pot fi apreciate ca având diferite intensităţi (ex: febră ridicată/ moderată/ uşoară sau parametri exprimabili numeric)

- ponderea simptomelor în aprecierea unui diagnostic nu este egală (de ex: concentraţia scăzută de hemoglobină este un indicator mai puternic pentru anemie decât paloarea)
- nu toate simptomele apar cu necesitate; există numeroase forme atipice sau asimptomatice
- metodele logice nu iau în considerare "prevalenţa" afecţiunilor (acest aspect deseori nu reprezintă un dezavantaj ci chiar un avantaj, sugerând şi posibilitatea unei afecţiuni mai rare).

Corectarea acestor neajunsuri a impus dezvoltarea unor metode ce depăşesc cadrul logicii bivalente şi care vor fi prezentate în continuare.

3. METODE STATISTICE. REGULA LUI BAYES

3.1. ASPECTE STATISTICE ÎN RAŢIONAMENTUL MEDICAL

O caracteristică esenţială a domeniului medical este faptul că majoritatea aspectelor întâlnite au un aspect probabilistic: rareori putem folosi termenii "întotdeauna" sau "niciodată", dar ne întâlnim frecvent cu termenii "adesea", "uneori", "rareori". Expresia probabilistică a terminologiei medicale este uneori acoperirea lacunelor cunoştinţelor disponibile care fac deocamdată imposibilă dihotomia necesară înlăturării unor ambiguităţi, dar trebuie să fim permanent conştienţi de caracterul statistic intrinsec al multor fenomene biologice, caracterul statistic fiind prezent încă de la nivel molecular, aşa că ecoul la nivel macroscopic nu este surprinzător. Abordarea probabilistică în medicină este un mod de gândire impus încă din faza de formare a viitorului medic, chiar dacă deocamdată este limitat în bună parte la utilizarea terminologiei "vagi" şi mai puţin la gândirea statistică propriu-zisă. Faptul că aspectul probabilist este fundamental în gândirea medicală a determinat dezvoltarea metodelor statistice de decizie medicală asistată, fiind nucleul unor metode care au depăşit cadrul statistic propriu-zis.

Abordarea statistică poate fi recomandabilă şi în situaţii în care caracterul intrinsec al aspectului nu este statistic, cum ar fi exprimarea intensităţii unui simptom sau ponderea simptomului pentru un anumit diagnostic.

Abordările statistice s-au orientat în două direcţii principale:

- regula lui Bayes
- pattern recognition.

3.2. REGULA LUI BAYES

Punctul de plecare al metodelor numite "bayesiene" îl constituie faptul că prezenţa unui simptom într-o boală poate fi dat cu o anumită probabilitate p(s+/b+), element fundamental deosebit de metodele logice în care simptomul putea fi doar prezent sau absent.

De asemenea, un simptom poate fi prezent în mai multe boli, deci se poate stabili o probabilitate de a găsi simptomul într-o populaţie, p(s+), indiferent de bolile cu care se asociază. Nu trebuie să uităm şi faptul că frecvenţa bolilor într-o populaţie poate fi caracterizată printr-o probabilitate p(b+). Deci întâlnirea unui simptom trebuie să ne sugereze posibilitatea unei boli cu o probabilitate p(b+/s+) dependentă de probabilităţile enumerate anterior.

Putem calcula ponderea simptomului s pentru diagnosticul bolii b după formula:

$$p(b+/s+) = p(s+/b+) \times p(b+)/p(s+) \qquad \text{(IV.1)}$$

cunoscută sub numele de "regula lui Bayes". Probabilităţile din membrul drept se consideră cunoscute pentru un anumit teritoriu într-un anumit moment. Ele ar putea fi cunoscute

printr-o analiză populațională (screening) sau prin evaluarea datelor de morbiditate pentru p(b+) şi din aprecierile unor experţi pentru p(s+/b+) şi p(s+). Literatura de specialitate apreciază că deocamdată ultimele două probabilităţi, în special p(s+), sunt cunoscute destul de aproximativ, limitând sau îngreunând aplicabilitatea metodei.

Tabel IV. 2. Tabel ilustrând frecvenţele cu care poate să apară într-o populaţie o anumită boală b şi relaţia ei cu simptomul s; tabele similare trebuie concepute pentru orice pereche b-s. În dreapta sunt prezentate ca exemplu probabilităţile necondiţionate p(s+) şi p(b+) şi condiţionate p(s+/b+), din care se va calcula p(b+/s+)

	s +	s -	
b +	**a**	**b**	**L1**
b -	**c**	**d**	**L2**
	C1	**C2**	**N**

p(s+) = C1/N
p(b+) = L1/N
p(s+/b+) = a/L1
p(b+/s+)=a/C1

De exemplu, să luăm o situaţie ce ar putea apărea la analiza fişelor unui eşantion de 4000 de indivizi. Dorim realizarea unui tabel ca tabelul IV.3 pentru o anumită boală (să zicem b = "viroză") şi un anumit simptom (să zicem s = "febră"). Din cei 4000 indivizi, în cursul anului precedent s-a semnalat prezenţa virozei la 100 de persoane; dintre acestea numai 80 au prezentat febră. De asemenea mai putem găsi că febra a fost semnalată şi la alte 70 de persoane care nu au avut viroză (dar poate au avut pneumonie, bronşită etc.). Trebuie subliniat că cele 3900 persoane care nu au avut viroză nu sunt presupuse sănătoase, cuprinzând atât persoane sănătoase cât şi persoane cu alte diagnostice; în tabelul nostru contează aici numai proprietatea "nu au avut viroză (b-)".

Tabel IV.3 Exemplu pentru ilustrarea regulii lui Bayes

	s +	s -	
b +	**80**	**20**	**100**
b -	**70**	**3850**	**3900**
	150	**3850**	**4000**

Reluând cele expuse: în total au avut viroză 100 de persoane din 4000, adică p(b+) = 100/4000; au avut febră 150 din 4000, adică p (s+) = 150/4000 etc. Aceste probabilităţi (p(b+), p(b-), p(s+), p(s-)) se numesc probabilităţi necondiţionate, toate acestea având la numitor numărul total al indivizilor (N=4000).

În cazul în care facem referire numai la o parte dintre ei (de ex. dintre cei care au avut viroză), atunci obţinem probabilităţi condiţionate. Iată un exemplu: din cei 100 care au avut viroză, 80 au prezentat febră, adică p(s+/b+) = 80/100; similar putem scrie orice probabilitate condiţionată; alt exemplu: din cei 3850 care nu au febră, 20 au avut viroză, adică p(b+/s-) = 20/3850. De obicei, o probabilitate condiţionată, de exemplu (p(b+/s-), se citeşte "probabilitatea ca un individ să aibe viroză (b+) dacă el nu are febră (/s-)".

Când avem la dispoziţie un astfel de tabel putem calcula direct p(b+/s+), dar de obicei astfel de tabele nu sunt disponibile. De aceea folosim regula lui Bazes în care înlocuim p(b+), p(s+) şi p(s+/b+) considerate din estimări în funcţie de acestea două (p(s+)). Putem astfel calcula: p(b+/s+) = 80/150. Această probabilitate reprezintă pentru programele de diagnostic asistat un element important deoarece ea arată cât de mult

contribuie faptul că un individ prezintă "febră" pentru a înclina spre diagnosticul de "viroză".

Pentru a stabili probabilitatea de a avea o anumită boală, trebuie să construim câte un tabel de genul tabelului III.3 pentru fiecare simptom; de fapt se construiesc tabele doar pentru simptomele caracteristice deşi actualmente dispunem de un număr relativ scăzut de date concrete.

Principial, termenii sunt accesibili, astfel încât îi putem considera pentru moment cunoscuți. Dacă am dispune de baze de date suficient de bogate, aceşti termeni ar putea fi calculați. Avem deci o evaluare a ponderilor simptomelor cu care am putea estima probabilitatea unui anumit diagnostic; câteva simptome cu ponderi ridicate ar conduce la un diagnostic foarte probabil. Din păcate, se mai impune o regulă: pentru ca aceste ponderi să fie combinate pentru mai multe simptome, este necesar ca simptomele să fie independente; ori, este cunoscut faptul că, foarte frecvent simptomele sunt corelate nu întâmplător, ci printr-o interdependență cauzală; utilizarea simultană a unor simptome puternic corelate conduce la creşterea artificială a probabilității unui diagnostic. Pentru evitarea acestor situații se pot aplica teste de independență (χ^2 = pătrat) din care se pot elimina simptomele redondante.

Cu toate dificultățile sale, regula lui Bayes a impus un punct de vedere mai realist, introducând şi o disciplină în elaborarea metodelor de diagnostic asistat.

4. PATTERN RECOGNITION

4.1. PRINCIPIUL METODEI "PATERN RECOGNITION"

Am păstrat denumirea originală a metodei deoarece traducerea uzuală ca "recunoaşterea formelor" este nepotriviti, mai ales în contextul diagnosticului asistat. Termenul de "pattern" trebuie înțeles ca un set de atribute caracteristice, neavând aici sensul de formă. Deşi este o metodă cu caracter statistic, principiul aplicării sale este fundamental diferit de cel al regulii lui Bayes.

Metoda "pattern recognition" face un pas însemnat spre apropierea de raționamentul medicului. Suntem de fapt obişnuiți cu operațiunea de recunoştere pe care o utilizăm frecvent; recunoaştem o persoană într-o fotografie, recunoaştem o voce, un mers, un obiect, o stare de spirit a unei persoane etc. Pe ce ne bazăm în aceste recunoaşteri? Analiza procesului de recunoaştere stă la baza metodei pe care o descriem în continuare.

Orice sistem sau obiect poate fi caracterizat printr-o imensitate de caracteristici exprimabile cantitativ sau calitativ care formează o mulțime a proprietăților, M. De exemplu, pentru o persoană această mulțime ar cuprinde: înălțimea, greutatea, vârsta, sexul, frecvența cardiacă, tensiunea sistolică, tensiunea diastolică, glicemia, proteinemia, proteinuria, pH-ul sanguin, numărul de respirații pe minut, indicele de memorie etc., etc. Unele mărimi nu se modifică (sexul, înălțimea la adult), altele se modifică lent (greutatea), altele au variații dependente de starea organismului (frecvența cardiacă).

Deci valorile proprietăților trebuie privite dinamic.

Indivizii (sau obiectele) unei populații pot fi împărțiți în clase după diferite criterii; de exmplu: după sex se împart în bărbați şi femei; după vârstă: copii şi adulți fiind posibil a diviza fiecare clasă în subclase (ex: adulți tineri, maturi şi vârstnici). Criteriile de includere a unui obiect într-o clasă sunt convenționale; totuşi, foarte multe "convenții" au intrat în viața de zi cu zi astfel încât de cele mai multe ori putem realiza clasificări fără a cunoaşte exact valorile numerice ale parametrilor implicați. Clasificările pomenite au apelat la câte o singură proprietate, sexul, respectiv vârsta.

Este cazul să facem două observații:

- prima: din larga paleta de caracetristici, o bună parte nu au putere de discriminare, având valori variind pe plaje similare pentru toate clasele (de ex: glicemia nu are putere de discriminare pentru stabilirea sexului sau a vârstei; ea devine însă o proprietate importantă la definirea clasei diabeticilor).

- a doua observație: indivizii din clase diferite se deosebesc prin mai multe caracteristici (de ex: valorile "normale" la femei sunt diferite de cele la bărbați pentru destul de mulți parametri; la fel, putem urmări variația unor valori "normale" cu vârsta); totuşi pentru definirea unei clase nu apelăm la toate deosebirile, mulțumindu-ne cu un număr redus de parametri, destul de des chiar cu unul singur, cum au fost clasificările date ca exemplu până acum.

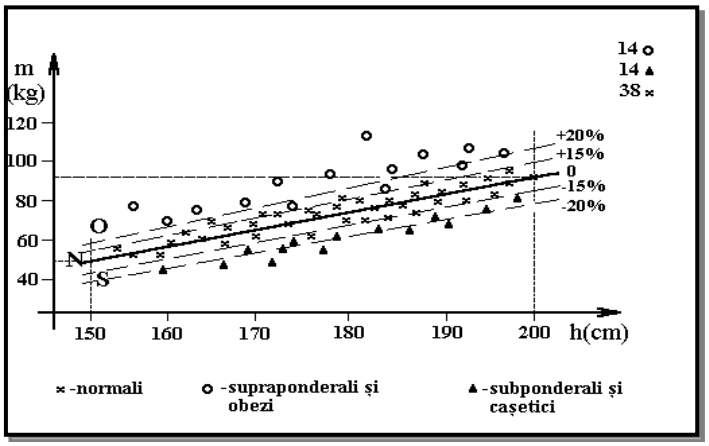


Figura IV.2. Reprezentarea în spațiul greutate – înălțime

Există însă situații în care clasele nu pot fi definite după un singur parametru; de exemplu putem împărți indivizii în: "obezi", "normali" şi "caşectici". Observăm imediat că greutatea nu este un parametru suficient; un individ cu greutatea de 75 kg pare obez la înălțimea de 1,55 şi slab la 1,95. Trebuie deci să definim clasele luând în considerare şi înălțimea. În asemenea situații se obişnuieşte să se facă reprezentări într-un sistem de coordonate având pe axe cei doi parametri numit spațiul stărilor. Într-un astfel de spațiu un individ este reprezentat printr-un punct. (fig.IV.2.)

Aprecierea apartenței la o clasă sau alta se poate face pe diferite criterii. În exemplul acesta apartenența la o clasă după aspectul fizic conduce la o clasificare cu sensibilitate şi specificitate destul de bună, deşi din punct de vedere medical este considerată nesatisfăcătoare. Când criteriul de clasificare nu utilizează pentru definirea claselor valorile parametrilor folosiți, spunem că avem un **clasificator independent**; dacă depinde exclusv de parametrii utilizați atunci avem un **clasificator formal**, iar dacă foloseşte numai unii din parametrii de reprezentare, eventual împreună cu alți parametri (adesea calificativi), atunci îl numim **clasificator parțial dependent.**

În cazul clasificatorilor formali împărțirea în clase este definită prin relații, cel mai adesea empirice, limitele dintre clase fiind convenționale şi deci oferind subiect de dispute. Este şi cazul exemplului dat în care domeniul greutăților acceptabile se defineşte în funcție de greutatea normală (sau ideală), care este dată de o relație de genul (IV.2):

$$m_i = (h-100) - (h-159)/4, \quad [h \text{ în cm}, m \text{ în kg}] \qquad \text{(IV.2)}$$

În jurul acestei drepte, într-un interval cu lățimea de 15% se consideră greutăți acceptabile, cei cu greutăți mai mari decât $1{,}15\ m_i$ fiind numiți *supraponderali*, iar dacă depăşesc

greutatea ideală cu peste 20% sunt numiți *obezi*; similar se definesc clasele de *subponderali*, respectiv *caşectici.*

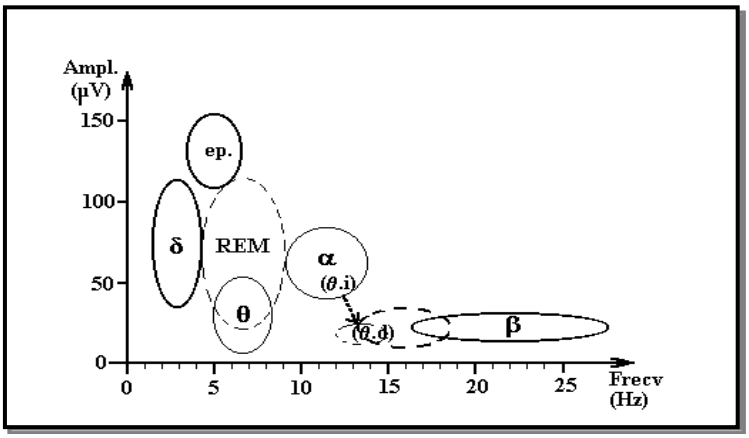


Figura IV.3. Caracterizarea undelor EEG într-un spațiu bidimensional: frecvență - amplitudine. Sunt ilustrate zonele undelor tipice α (o.i. = ochii închiși, od = ochii deschiși), β, σ, δ, complexele vârf-undă din descărcări epileptice (ep.) şi undele din faza de somn "Rapid-Eye-Movement" (REM)

Cele mai interesante reprezentări apar în cazul clasificatorilor independenți sau parțial dependenți, situații în care delimitarea dintre clase nu mai este bine deifintă şi în practică putem întâlni şi zone de intersecție (mulțimile nu sunt disjuncte). În fig. III.3 am reprezentat undele EEG tipice în spațiul amplitudine-frecvență cu un clasificator parțial dependent. Ne putem imagina şi o reprezentare cu un clasificator independent, de exemplu luând 3 loturi: un lot de sportivi, unul de indivizi sănătoşi nesportivi şi un lot de astmatici şi să-i reprezentăm într-un sistem de coordonate, să zicem capacitate vitală - VEMS (volum expirator maxim pe secundă). Clasele se vor suprapune parțial însă pentru fiecare grup va exista zona sa de dominanță. În metoda "pattern-recognition" aplicată diagnosticului asistat clasificatorul este cel mai adesea parțial dependent sau independent, mulți dintre parametrii utilizați fiind calitativi, deci fără posibilitatea de a fi plasați pe axe.

Problema principală care se pune este ca, având la dispoziție ansamblul de date, cantitative şi calitative atât pentru loturi de pacienți grupați pe diagnostice cât şi pentru loturi de control (normali), să găsim cel mai potrivit sistem de axe de coordonate care ne permite să reprezentăm loturile prin regiuni cât mai compacte şi, dacă este posbil, ca acestea să fie disjuncte; această fază se numeşte "selecția atributelor". După ce s-au găsit parametrii potriviți se calculează domeniile lor de variație pentru fiecare clasă, definindu-se regiunile din spațiu ale claselor, etapă numită "învățare supervizată". În această fază calculatorul este aproape gata să stabilească diagnosticul unui pacient reprezentându-l în spațiu şi găsind în ce domeniu, din cele definite anterior, se găseşte. Un ultim aspect care ar mai trebui rezolvat se referă la criteriul după care luăm decizia de apartenență la o clasă, deci mai este necesară etapa de definire a unei funcții de decizie. Toate aceste etape au un suport matematic solid şi au fost prezentate sumar în capitolul de prelucrare a semnalelor biologice. Acum nu ne oprim în detaliu la aceste aspecte, trecând doar rapid în revistă elementele esențiale ale fiecărei etape.

4.2. ETAPELE APLICĂRII METODEI "PATTERN RECOGNITION". CLASIFICAREA METODELOR

a) Alegerea atributelor ("feature selection")

Cel mai important aspect în recunoaşterea unui pattern însă şi cel mai mult discutat este cel referitor la alegerea atributelor cu putere reală de descriminare a claselor. Din imensitatea de mărimi caracteristice ale unui obiect sau individ (R), reținerea celor mai "importante" caracteristici poate fi legată de îndeplinirea unor condiții pentru fiecare atribut selectat:

- să aibe o variabilitate mică în interiorul claselor (minimizarea "distanțelor" intraclasă)
- să ofere o discriminare satisfăcătoare între clase (maximizarea "distanțelor" interclase).

S-au propus o serie de metode pentru optimizarea selectării aributelor, bazate pe transformări de coordonate, fie rotații, fie introducerea unor ponderi. Se încearcă şi găsirea unor noi variabile prin "combinarea" caracteristicilor inițiale, însă numărul acestor combinații este tot foarte mare. Rezultatele care se obțin sunt satisfăcătoare pentru aplicarea practică, matematicienii fiind mai curând nemulțumiți nu de rezultate ci de absența unui criteriu nediscutabil de optimizare. Trebuie să mai menționăm că, după depistarea caracteristicilor eligibile, se verifică şi gradul de corelație între ele; în cazul unor corelații puternice între două caracteristici se păstrează numai una, informațiile aduse de cealaltă fiind redondante. Se ajunge în final la selectarea unui număr deosebit de scăzut de atribute (N<<R) care păstrează (uneori chiar îmbunătățesc) capacitatea de discriminare între clase.

b) Învățarea supervizată. Definirea criteriului de clasificare

În faza de alegere a atributelor am ajuns la precizarea în spațiul cu N dimensiuni a poziției fiecăreia din cele K clase; spunem că am realizat o "învățare" supervizată - la fel cum fiecare persoană de fapt trebuie întâi să învețe pentru a putea apoi realiza o recunoaştere. Să presupunem că împărțirea în clase a fost realizată conform graficului din fig.IV.2, separarea între clasele notate O, N şi S fiind dată de liniile de 15%. Introducând acum datele unui individ despre care nu ştim în ce clasă face parte, el va aparține unei regiuni deci poate uşor fi inclus într-o clasă şi - important - putem face clasificarea chiar fără a mai efectua celelalte calcule. Principii similare stau la baza construcției nomogramelor dacă numărul de variabile este scăzut. Trebuie menționat totuși că acest gen de clasificare este posibil numai dacă ansamblul claselor umplu întregul spațiu, deci când separarea între clase se face prin plane (sau suprafețe curbe) în spațiul cu N dimensiuni (hiperplane).

Există cazuri în care regiunea unei clase nu este definită printr-o relație ce ne permite extinderea sa, ea fiind restrânsă la spațiul delimitat de valorile din setul de învățare (ca de ex. în fig.IV.3); astfel de situații apar mai ales când criteriul de clasificare este extern - independent sau parțial dependent. În aceste situații este posibil să întâlnim obiecte (indivizi) care nu aparțin nici unei clase întâlnite anterior. Încadrarea sa într-o clasă se poate face în funcție de "apropierea" de una din clase, deci trebuie să alegem un criteriu geometric şi să măsurăm distanța de la obiectul nostru până la clasele definite (s-au propus variante de distanțe: până la "centrul" clasei sau până la suprafața clasei).

În măsurarea distanțelor într-un spațiu multidimensional apare o problemă deosebită: mărimile de pe axe au - aproape întotdeauna - unități de măsură diferite, fiind deci incomparabile! O soluție este "normalizarea" lor obținută cel mai adesea prin transformarea tuturor mărimilor în procente: se alege pentru fiecare mărime o referință (fie valoarea maximă, fie cea medie, fie derivația standard etc.) şi se transformă toate datele în procente față de referință. Această operațiune se face de fapt înainte de selecția atributelor, fiind utilă şi în faza respectivă.

Deci într-un spațiu normalizat putem compara distanțele şi vom putea stabili un criteriu după care să clasificăm orice obiect când sunt cunoscute caracteristicile claselor. Reguli de acest gen se pot aplica şi în cazul claselor definite prin clasificatori formali.

c) Învăţare nesupervizată. Metoda grupării

Sistemul descris anterior, de învăţare supervizată, presupune o cunoaştere apriori a claselor, deci în faza de învăţare se întroduc în calculator, pe lângă toate mărimile culese despre obiect, şi informaţii privind clasa căreia îi aparţine. În aceste situaţii calculatorul este folosit numai pentru a categorisi obiecte introduse ulterior în calculator în aceste clase.

Avem însă posibilitatea de a urmări modul în care sunt distribuite punctele, care reprezintă obiectele (indivizii), într-un spaţiu multidimensional fără a defini de la început nici un fel de clase. În cazul în care găsim regiuni de concentrare a punctelor ("clusters"), putem defini clase în jurul lor. Spunem în acest caz că avem o învăţare nesupervizată, clasele nefiind definite apriori ci fiind "descoperite" de calculator.

Metoda "pattern recognition", deşi are unele calităţi prin care depăşeşte substanţial alte metode de diagnostic asistat, nu are o răspândire prea largă datorată numărului mare de parametri ce trebuie culeşi în faza iniţială precum şi numărului foarte mare de calcule necesar pentru optimizări şi decizii, ceea ce a impus de la început utilizarea unor calculatoare performante, accesibile doar în centrele mai importante de cercetare. Se poate remarca totuşi o creştere a numărului de aplicaţii în ultimul timp, în ciuda concurenţei puternice a metodei ce o vom descrie în continuare - sitemele expert.

5. ELEMENTE DE LOGICĂ

Metodele expuse până acum, cu performanţe mai ridicate sau mai modeste, suferă toate de un neajuns: utilizează un format de exprimare diferit de cel curent al medicilor. În practica uzuală informaţia se transmite prin propoziţii, iar raţionamentul medical se construieşte prin operaţii cu aceste propoziţii. Pare deci cât se poate de natural ca cele mai agreate metode de diagnostic asistat s-au dovedit cele pe care le numim "euristice" (l. gr. *heuriskein* - a descoperi) în care poziţia centrală o ocupă *sistemele expert*. În aceste metode atât baza de cunoştinţe cât şi vectorul de stare al pacientului se exprimă prin propoziţii, cu care se efectuează operaţii logice. Vom prezenta în continuare sintetic câteva noţiuni fundamentale de logică.

5.1. NOŢIUNI GENERALE

a) Propoziţia

În viaţa curentă ideile noastre se transmit prin propoziţii ce exprimă proprietăţile unor obiecte, cauzele unor evenimente, sau exprimă întrebări, dorinţe, porunci etc. Putem deci clasifica propoziţiile în cel puţin 3 clase mari:

- propoziţii cognitive = însuşiri ale obiectelor, cauzele unor evenimente
- propoziţii interogative = întrebări
- propoziţii imperative = ordine, dorinţe.

Ex. Asmul bronşic este o boală a aparatului respirator este o propoziţie cognitivă..

Numai propoziţiile cognitive au asociată o valoare de adevăr, adică o propoziţie poate fi:

- adevărată (A sau 1)
- falsă (F sau 0)
- nesigură (?).

b) Forma logică

Exprimarea ideilor se realizează în cadrul unor scheme, cu diferite grade de complexitate pe care le numim forme logice. Am întâlnit deja o formă logică: propoziţia. O propoziţie este de fapt aplicarea unei operaţii logice (de ex. afirmaţia) asupra unor forme

logice mai simple, numite noţiuni. În exemplul de mai sus am utilizat două noţiuni: *asmul bronşic,* respectiv *boală a aparatului respirator*. Prima (asmul bronşic) reprezintă obiectul gândirii şi în propoziţie se va numi subiect logic, notat cu S iar a doua noţiune (boală a aparatului respirator) redă "ce se spune despre subiect" şi se va numi predicat logic, notat cu P. Operaţia care leagă aici S şi P este afirmativă şi putem nota formula generală a unei propoziţii afirmative:

S este **P**

Pentru o propoziţie negativă formula generală este:

S nu este **P**

Ex. Astmul bronşic nu este boală infecţioasă

Pe lângă formele logice simple menţionate deja - noţiunea şi propoziţia - există şi o formă logică mai complexă, inferenţa, care cuprinde mai multe propoziţii; unele numite premise din care construim o propoziţie derivată numită concluzie. Ex.:

Premise: - Hemoragia duce la scăderea masei eritrocitare
- Eritrocitele conţin hemoglobină

Concluzie: - Hemoragia duce la scăderea hemoglobinei (anemie)

Observăm deci că formele logice pot fi ierarhizate: noţiunea, propoziţia, inferenţa.

c) Propoziţii categorice

Propoziţiile categorice sunt cele mai simple propoziţii logice,exprimând un singur raport între două noţiuni, fără nici o condiţie. Ele pot fi:

- universale:
  - = afirmative: Toţi **S** sunt **P**
  - = negative: Nici un **S** nu este **P**
- particulare:
  - = afirmative: Unii **S** sunt **P**
  - = negative: Unii **S** nu sunt **P**
- singulare:
  - = afirmative: Acest **S** este **P**
  - = negative: Acest **S** nu este **P** (IV.3)

În logică propoziţiile singulare sunt testate ca universale, **S** fiind o clasă cu un singur element.

d) Pricipiile logicii

Aplicarea unor operaţii asupra formelor logice pentru a obţine propoziţii noi trebuie să respecte o serie de legi de raţionare, dintre care patru au un caracter fundamental şi se numesc principii logice.

- Principiul identităţii: un obiect este inconfundabil cu alt obiect.
- Principiul non-contradicţiei: o propoziţie nu poate fi şi adevărată şi falsă în acelaşi timp.

- Principiul terţului exclus: o propoziţie - într-un context - poate fi fie acceptată, fie neacceptată; nu trebuie confundat cu principiul bivalenţei privind valoarea de adevăr a propoziţiei (stă la baza demonstraţiei prin reducere la absurd).
- Principiul raţiunii suficiente: nici o propoziţie nu este acceptată sau respinsă într-un raţionament decât dacă există o justificare pentru acceptare (respingere, necesitate şi suficienţă).

5.2. PROPOZIŢII COMPUSE

Prin aplicarea unor operaţii logice asupra unor propoziţii simple se obţine o formă logică nouă - propoziţia compusă. Valoarea de adevăr a propoziţiei compuse este dependentă de valoarea de adevăr ale propoziţiilor simple şi este redată uzual sub forma unor "tabele de adevăr", în care se notează valoarea "adevărat" cu 1 şi "fals" cu 0. Prezentăm în continuare operaţiile logice posibile pentru construcţia unor propoziţii compuse, împreună cu tabele de adevăr asociate.

a) Negaţia: notată ' ~ ' sau ' ¬ '; este un operator ' unar' ~ p se citeşte 'non-p'; se mai numeşte 'NOT'

q	1	0
~ p	0	1

b) Conjuncţia: notată 'Λ' sau '&'; este un operator binar; p & q se citeşte 'p şi q'; se mai numeşte 'AND'

p / q	1	0
1	1	0
0	0	0

c) Disjuncţia: notată 'V'; pVq se citeşte 'p sau q'; se mai numeşte 'OR'

p / q	1	0
1	1	1
0	1	0

d) Disjuncţia exclusivă: notată ' ⊗ ' sau 'W'; pWq sw citeşte 'sau p, sau q'; se mai numeşte 'XOR'

p / q	1	0
1	0	1
0	1	0

e) Implicaţia: notată ' $\rightarrow$ '; p $\rightarrow$ q se citeşte 'dacă p atunci q'; p se numeşte antecedent, q se numeşte consecvent

p / q	1	0
1	1	1
0	0	1

f) Echivalența : notată ' $\leftrightarrow$ ' sau ' $\equiv$ '; $p \equiv q$ se citeşte ' dacă şi numai dacă p atunci q'

p / q	1	0
1	1	0
0	0	1

5.3. INFERENȚE LOGICE

O formă logică mai complexă decât propoziția este inferența, prin care, din unele propoziții (premise) se construiește o propoziție nouă (concluzie).

Există două clase mari de inferențe:

- deductive - de la general spre particular
- inductive - de la particular spre general.

Sistemele expert actuale utilizate în domeniul medical utilizează exclusiv inferențe deductive. Vom prezenta în continuare prtincipalele operații cu propoziții compuse. Vom nota inferența cu semnul ' $\vdash$ ', plasând premisele în stânga şi concluzia în dreapta.

a) Modus ponens:

$$\begin{array}{l} p \rightarrow q \\ p \end{array} \vdash q$$

b) Modus tollens:

$$\begin{array}{l} p \rightarrow q \\ \sim q \end{array} \vdash \sim p$$

c) Silogismul:

$$\begin{array}{l} p \rightarrow q \\ q \rightarrow r \end{array} \vdash p \rightarrow r$$

Ar putea fi desigur analizate detaliat cazurile ce apar în funcție de tipul fiecărei propoziții p, q, r, însă nu ne-am propus prezentarea acestor detalii, scopul principal fiind reamintirea elementelor fundamentale de logică ce stau la baza construcției unora dintre cele mai performante programe de informatică medicală: sistemele expert.

5.4. ELEMENTE ALE LIMBAJULUI PROLOG

Majoritatea limbajelor de calculator sunt orientate pentru rezolvarea unor probleme numerice, având ca operatori fundamentali operațiile aritmetice. Lucrul cu propoziții necesită însă un limbaj adecvat, în care operatorii fundamentali să fie operatorii logici "nu, şi, sau". Un astfel de limbaj este 'PROLOG' (PROgramming in LOGics), dedicat pentru transpunerea formală a propozițiilor şi operații cu propoziții. Rularea unui program implică estimarea valorii de adevăr a fiecărei propoziții, iar rezultatul este întotdeauna fie o valoare de adevăr, fie o propoziție.

Elementul fundamentale ale limbajului PROLOG sunt:

- **predicate:** care exprimă o relație între obiecte, de obicei primul având rolul subiectului dintr-o propoziție logică iar al doilea, eventual şi celelalte fiind obiect al acțiunii sau proprietăți
- **clauze:** reprezintă fapte sau reguli; ele se construiesc folosind predicatele enumerate în secțiunea de predicate având înlocuite concret valori pentru parametrii de paranteză
- **domenii:** reprezintă paragraful de început al unui program în PROLOG în care sunt enumerate tipurile de variabile / parametri întâlniți în clauze.

Exemplu:

```
domains
        diagn           = symbol
        tens_art_sis    = integer
        hemoglobin      = real
predicates
        are (pacient, diagn, tens_art_sis, hemoglobin)
clauses
        are (X, anemie, _ , Y) if  Y < 11.5
        are (X, hipertensiune, Z, _) if  Z >150
        are (X, sănătos, T,Y) if  Z <= 150 and  Y >= 11.5)
        are (popescu, Diag, 12, 140)
        are (ionescu, Diag, 11, 135)
        are (petrescu, Diag, 13, 160)
        are (vasile, Diag, 10, 155).
```

Observăm că în clauze apar două feluri de propoziții: primele trei reprezintă *reguli*, (conținând variabile) iar următoarele 4 sunt *fapte*, conținând date.

Programul solicită operatorului rezolvarea unei probleme ('goal'); de ex. pentru un caz particular:

```
goal:
are (ionescu, Diagn, 11, 135)
```

vom primi răspunsul:

```
Diagn = anemie
```

iar pentru un caz general:

```
goal:
are (X, Diagn, _ , _)
```

vom primi răspunsul:

```
X = popescu, diagn, = sănătos
X = ionescu, diagn = anemie
X = petrescu, diagn = hipertensiune
```

X = vasile, diagn = anemie
X = vasile, diagn = hipertensiune
5 solutions.

Desigur în situaţii reale avem un grad mult mai ridicat de complexitate. Exemplele prezentate sunt foarte simple, însă ilustrează modul în care se formalizează informaţiile în cadrul sistemelor expert.

6. SISTEME EXPERT

6.1. STRUCTURA UNUI SISTEM EXPERT

În figura IV.4 este prezentată schematic structura unui sistem expert.

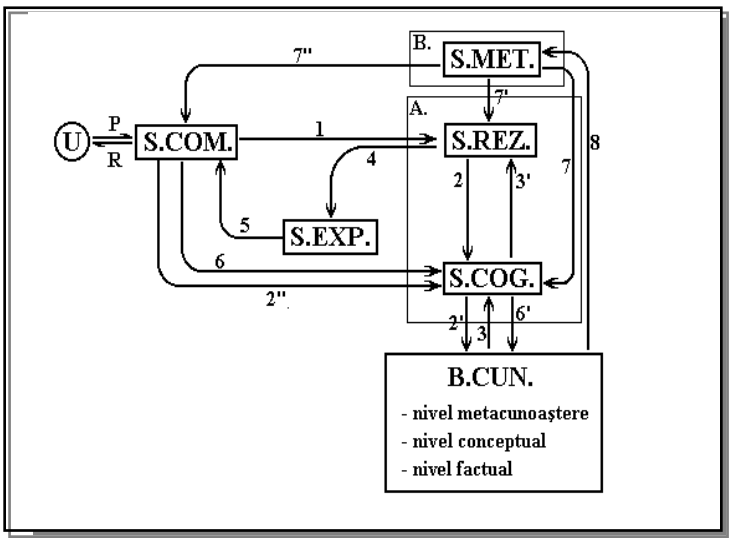


Figura IV.4. Structura unui sistem expert

Să descriem pe scurt sistemele componente.

a) **B.CUN** - baza de cunoştinţe: este elementul fundamental al unui sistem expert în care cunoştinţele sunt de obicei clasificate pe trei nivele:

- nivelul cunoştinţelor factuale, care cuprinde fapte reprezentate ca instanţe ale conceptelor (cunoştinţe empirice)
- nivelul cunoştinţelor conceptuale, care cuprinde cunoştinţe teoretice, reliefând legăturile şi relaţiile cauzale între elemente
- nivelul de metacunoaştere, care cuprinde cunoştinţele despre cunoaştere şi reprezentările sale.

Metodele de realizare a bazelor de cunoştinţe constituie un dezvoltat capitol al inteligenţei artificiale numit "ingineria cunoaşterii" (knowledge engineering) sau "reprezentarea cunoştinţelor" (knowledge representation). Cele mai utilizate metode de reprezentare a cunoaşterii sunt:

- reprezentarea prin formalizare în calculul predicatelor
- metode procedurale
- reţele semantice
- sisteme de producţie ("production rules")

- reprezentare prin cadre
- reprezentare cu hiper-rețele.

Nu ne oprim aici pentru descrierea lor; menționăm doar că van Bemmel semnala încă în 1985 că formalizarea cunoştinţelor medicale, care are deocamdată un nivel deosebit de scăzut, necesită o pregătire teoretică adecvată, insistând asupra introducerii biomatematicii în programa facultăților de medicină.

b) **S.COG.** - sistemul cognitiv: asigură accesul la baza de cunoştinţe şi are în principal două sarcini:

- căutarea pieselor de cunoaştere în baza de cunoştinţe (fie prin simboluri fie prin proprietăți)
- crearea şi actualizarea bazei de cunoştinţe (prin adăugiri, ştergeri sau modificări).

c) **S.REZ.** - sistemul rezolutiv (l.engl: inference machine) este modulul central din program ce are ca obiectiv rezolvarea problemelor puse de utilizator; în funcție de gradul de complexitate el poate realiza:

- alegerea strategiei de control adecvate problemei
- elaborarea planului de rezolvare
- desfăşurarea acțiunilor din plan
- trasarea drumurilor de raționament prin arborii deductivi
- constituirea informației de control
- verificarea paşilor de rezolvare.

Performanţele unui sistem expert sunt determinate în principal de calitatea sistemului rezolutiv. Actualmente sunt realizate sisteme expert acoperind o gamă largă de performanţe, de la sisteme simple, în care sistemul rezolutiv se limitează la cererea pieselor de cunoaştere şi estimarea unor "potriviri" între situația reală şi diferite piese de cunoaştere, până la sisteme sofisticate, cu elaborare de strategii şi capabile să schimbe strategia de rezolvare dacă este cazul.

d) **S.EXP.** - sistemul explicativ: are ca sarcină principală justificarea soluțiilor oferite de sistemul expert la problemele puse, prin:

- listarea şi/sau interpretarea drumurilor de raționament ale sistemului rezolutiv
- editarea cauzelor greşelilor sau eşecului în găsirea unei soluții
- evidențierea pieselor de cunoaştere care lipsesc din lanțul inferențial.

Sistemul explicativ poate chiar lipsi sau poate fi foarte simplu (enumerarea pieselor de cunoaştere folosite), dar poate fi elaborat până la justificarea fiecărui pas din raționament sau evidențierea unor piese de cunoaştere contradictorii sau suspecte. Sistemul explicativ are un rol deosebit de important în utilizarea sistemelor expert în procesul didactic.

e) **S.COM.** - sistemul de comunicare: asigură interfața cu utilizatorul; deşi dialogul este dirijat de sistemul rezolutiv, sistemele de comunicare evoluate pot conține:

- procesoare pentru limbaje de reprezentare a cunoaşterii
- procesoare pentru achiziția semnalelor sau imaginilor, ieşiri grafice, conexiuni cu alte echipamente etc.

f) **S.MET.** - sistemul metarezolutiv: este inclus în schemele clasice ale sistemelor expert deşi sistemele realizate până în prezent încă nu îl conțin. Sarcina acestui sistem ar fi adecvarea şi validarea mecanismelor fundamentale utilizate de sistemul rezolutiv (sau cognitiv), evaluarea caracteristicilor domeniului de expertiză, preluând sarcini dirijate ale sistemului rezolutiv privind priorități şi restricții de aplicare ale strategiilor de rezolvare.

Sistemele expert medicale realizate şi aplicate până în prezent, deşi au diferite nivele de dezvoltare, au căutat în special dezvoltarea bazelor de cunoştinţe având sistemul rezolutiv şi cel explicativ la nivele relativ modeste comparativ cu nivelul teoretic în domeniu şi acesta s-ar datora într-o bună măsură nivelului încă nesatisfăcător al formalizării cunoştinţelor medicale. Aceasta este şi direcția în care se depun actualmente

cele mai mari eforturi, fiind necesară în paralel şi o pregătire corespunzătoare a potenţialilor utilizatori.

6.2. DESCRIEREA CONEXIUNILOR

Vom încerca să descriem într-o formă simplificată aspectele funcţionale fundamentale ale unui sistem expert prin descriere a conexiunilor între subsisteme.

Utilizatorul U formulează problema (întrebarea) notată în fig.3 prin conexiunea P. Sistemul de comunicaţie care asigură interfaţa cu ansamblul de prelucrare a cunoaşterii despre domeniu transformă problema P a utilizatorului în "problema bine definită" - 1, care este transmisă sistemului rezolutiv. Acesta alege o strategie, elaborează un plan de rezolvare şi, pentru realizarea acestui plan solicită piese de cunoaştere de la sistemul cognitiv (conexiunea 2), care o solicită bazei de cunoştinţe (conexiunea 2'); piesa de cunoaştere este transferată sistemului rezolutiv (conexiunile 3 şi 3'). Ciclul 2-2'-3-3' se repetă de câte ori este necesar până când sistemul rezolutiv fie găseşte soluţia (sau soluţiile) problemei, fie abandonează căutarea soluţiei din diferite motive, ce vor fi comunicate utilizatorului. Rezultatul obţinut este transmis prin conexiunea 4 către sistemul explicativ care transformă în formă inteligibilă mesajele sistemului rezolutiv, transmiţându-le prin sistemul de comunicaţie (conexiunea 5) către utilizator, sub forma răspunsului R. După comunicarea răspunsului, utilizatorul poate investiga mai în detaliu modul de rezolvare obţinând - la cerere - întregul traseu al raţionamentelor; se pot astfel evidenţia cauzele abandonului, piesele de cunoaştere lipsă sau contradictorii, date insuficiente în problemă etc.

Este evident că actualizarea bazei de cunoştinţe este strict necesară pentru obţinerea unor rezultate de încredere. Pentru această operaţie utilizatorul poate analiza oricând piesele de cunoaştere, direct prin conexiunile 2" şi 2', fără să fie necesară o problemă pentru a fi solicitate. Introducerea unor noi piese de cunoaştere se face tot direct, prin conexiunile 6 şi 6'.

Adăugarea unui sistem metarezolutiv ar permite supravegherea unor acţiuni în cursul funcţionării sistemului expert prin conexiunile 7, 7' şi 7" de adecvare a mecanismelor fundamentale. Sistemul metarezolutiv ar folosi cunoştinţele la nivelul metcunoaşterii din baza de cunoştinţe prin conexiunea 8.

6.3. CARACTERISTICILE PRINCIPALE ALE SISTEMELOR EXPERT

După ce peste două decenii progresele înregistrate în domeniul diagnosticului asistat au fost relativ modeste, apariţia sistemelor expert a revigorat speranţele în această direcţie. Ce au sistemele expert deosebit faţă de celelalte modele?

a) Sistemele expert au trecut de la reprezentarea empirică a cunoştinţelor la reprezentări adecvate pentru complexitatea lor, fiind clasificate pe nivele şi cuprinzând şi relaţiile.

b) Raţionamentele folosite de sistemele expert se fac prin mecanisme inferenţiale, depăşind simplele structuri liniare, folosind mecanisme de căutare adecvate ce permit creşterea deosebită a performanţelor.

c) Sistemele expert pot extrage cunoştinţe din baze de date.

d) Sistemele expert se încadrează în clasa programelor de inteligenţă artificială prin caracteristica lor de a putea învăţa.

Iată ca exemplu o secvenţă din rularea unui program. Din meniul principal se alege modul de lucru: "introducerea simptomelor"; o porţiune de dialog este prezentată mai jos:

...
> dureri?
: da
> localizare
: piept
> se agravează la efort?
: da
> palpitații?
: da
> hipertrofie ventriculară?
: da
> stângă/dreaptă?
: stânga
...

După introducerea simptomelor se revine în meniul principal şi se alege modul de lucru: "diagnostic"; ilustrăm din nou o secvență:

>> diagnostic propus: hipertensiune
>> sunteți de acord cu diagnosticul propus? (da/nu)
: nu
>> care este diagnosticul dvs.?
: angină pectorală

În această fază programul caută în baza de cunoştințe "angină pectorală" şi simptomele sale, făcând comparație cu simptomele introduse. Dacă "angina pectorală" există în baza de cunoştințe calculatorul va evidenția - din filmul rulării executate - întrebarea la care decizia sa a ales altă ramură, precizând:

>> nu am ales: angină pectorală
>> deoarece la întrebarea: "transpirație rece?"
>> ați răspuns: "nu"
...

Dacă în baza de cunoştințe nu figurează "angina pectorală", dialogul ar contiua astfel:

>> piesă de cunoaştere absentă în baza de cunoştințe: angina pectorală
> prin ce se deosebeşte : angina pectorală
> de: hipertensiune
> proprietate nouă #1:
: dispnee
> proprietate nouă #2:
: end
> doriți să salvăm noul diagnostic?
: da

...

Am introdus astfel o nouă piesă de cunoaştere care moşteneşte proprietățile de la precedenta şi în plus are o proprietate în plus, inexistentă la precedenta. Noua piesă împreună cu proprietățile sale va fi introdusă în baza de cunoştințe. La o nouă rulare a

programului, răspunzând în acelaşi mod ca în rularea precedentă, dialogul se va desfăşura la fel până când se propunea un diagnostic, însă acum va mai apare o întrebare:

> dispnee?
: da
...
>> diagnostic propus: angină pectorală
...

În exemplul dat am folosit o parte din baza de cunoştinţe a sistemului expert INTERNIST însă, la fel ca multe alte sisteme expert, pentru creşterea operativităţii în manevrare, simptomele nu se introduc prin dialog ci prin selecţie dintr-o listă de mari dimensiuni (de ex. pentru versiunea INTERNIST folosită de noi această listă de simptome se intinde pe 34 pagini-ecran).

Exemplul dat este simplificat faţă de programul real, care prezintă de fapt o listă de diagnostice ierarhizate după procentul de potriviri. De asemenea se iau măsuri şi pentru asigurarea că baza de cunoşinţe nu va fi modificată de persoane neautorizate. De obicei un sistem exepert de uz practic are un administrator care răspunde de integritatea bazei de cunoştinţe.

6.4. SISTEME EXPERT MEDICALE

Primul sistem expert, DENDRAL, a fost realizat în 1964 şi se referea la structurile moleculelor organice. În domeniul medical primul sistem expert, de mare succes, a fost MYCIN realizat de colectivul condus de Shortliffe pentru diagnostic în infecţii bacteriene ale sângelui, bazat pe simptome şi date de laborator; după estimarea diagnosticului sistemul făcea şi propuneri de tratament medicamentos. Concepţii similare au stat şi la baza altor două sisteme expert: HEADMED pentru patologia neuro-psihiatrică şi PUFF pentru boli pulmonare. Cel mai cunoscut sistem expert este INTERNIST, pentru asistarea diagnosticului în medicina internă; sistemul este intens folosit şi în scopuri didactice în multe universităţi din lume. Un alt sistem, expert într-un domeniu mai îngust, însă de mare utilitate practică este VM (Ventilator Monitor); acesta este capabil să supravegheze funcţionarea plămânului artificial în saloanele de terapie intensivă, să ia decizii în foarte multe situaţii şi să avertizeze, eventual, personalul în situaţii delicate. Sistemul expert CASNET, pentru diagnostic în boli de ochi este unul dintre sistemele expert cele mai bine elaborate prin organizarea bazei de cunoştinţe. Din punct de vedere al aplicaţiilor directe probabil că cel mai utilizat este sistemul TROPICAID, elaborat pentru asistarea diagnosticului în ţări tropicale, ce permite rezolvarea a peste trei sferturi din cazuri de către un cadru mediu, selectând astfel pentru asistenţa medicală mai elaborată numai cazurile mai deosebite. În ultimul timp au apărut sisteme expert "independente de domeniu", numite şi "shell" (de exemplu INTEXP) adică sisteme expert cu structură flexibilă în care utilizatorul îşi poate introduce baza proprie de cunoştinţe, obţinând astfel un sistem expert specializat.

Deşi eforturile depuse pentru realizarea sistemelor expert au fost foarte mari iar unele sisteme au ajuns destul de performante, aplicarea lor concretă este încă relativ limitată, dar în creştere constantă, odată cu creşterea dotării cu tehnică de calcul şi creşterea nivelului de pregătire al utilizatorilor. Sistemele expert vor constitui probabil un instrument omniprezent în clinici şi cabinete medicale, chiar dacă vor fi folosite numai pentru "asistarea" medicului în faza de diagnostic.

7. ESTIMAREA CALITĂȚII CLASIFICĂRII

Sistemele de diagnostic asistat realizează, principial, o operație de clasificare. Utilizarea calculatoarelor pentru asistarea în aceste operațiuni nu este lipsită de riscul unor clasificări greşite. De aceea, este important a avea criterii bine definite de apreciere a calității unui clasificator. (Menționăm aici că aceste criterii au un caracter general, nefiind restrânse ca aplicabilitate numai la diagnosticul asistat).

Aprecierea porneşte desigur de la confruntarea cu realitatea (Tabel IV.4.). Să considerăm că dintr-un total de N indivizi, L_1 sunt pozitivi (de exemplu au o boală) şi notăm cu L_2 restul indivizilor, care sunt negativi din punct de vedere al afecțiunii respective (atenție: nu impunem alte condiții, deci nu înseamnă că L_2 sunt sănătoşi - ei pot avea alte afecțiuni). Clasificatorul pe care dorim să-l analizăm - în cazul nostru va fi un algoritm de clasificare - face o clasificare corectă a n_{11} indivizi dintre cei L_1; aceştia se vor numi *'real pozitivi'* (R+). Restul subiecților, până la L_1, (adica n_{12}), au fost apreciați ca neaparținând clasei K - se vor numi în acest caz *'fals negativi'* (F-). Dintre cei L_2 care sunt negativi, un număr n_{22} au fost clasificați corect, ca neaparținând clasei K (*real negativi* R -), dar n_{21} au fost clasificați greşit, ca aparținând clasei K – aceştia sunt *'fals pozitivi'* (F +).

Tabel IV.4. Estimarea calității clasificatorului

	clasificator K+	K-	
B real	n_{11}	n_{12}	L_1
B-	n_{21}	n_{22}	L_2
	C_1	C_2	N

Se folosesc uzual trei perechi de indicatori pentru a estima calitatea clasificării.

a) Perechea sensibilitate (SN) – specificitate (SP)

Sensibilitatea reprezintă capacitatea clasificatorului de încadrare corectă a celor pozitivi, iar specificitatea este capacitatea de rejecție corectă a celor negativi:

$$SN = p(K+/B+) = n_{11} / L_1$$

$$SP = p(K-/B-) = n_{22} / L_2$$

b) Valoarea predictivă pozitivă (VPP) şi valoarea predictivă negativă (VPN)

Valoarea predictivă pozitivă (VPP) este definită prin proporția încadrării corecte a celor declarați pozitivi iar valoarea predictivă negativă (VPN) prin rata încadrării corecte a celor clasificați negativi:

$$VPP = p(B+/K+) = n_{11} / C_1$$

$$VPN = p(B-/K-) = n_{22} / C_2$$

c) Indicatori globali – acuratețea şi rata erorii de clasificare

$$AC = (n_{11} + n_{22}) / N$$

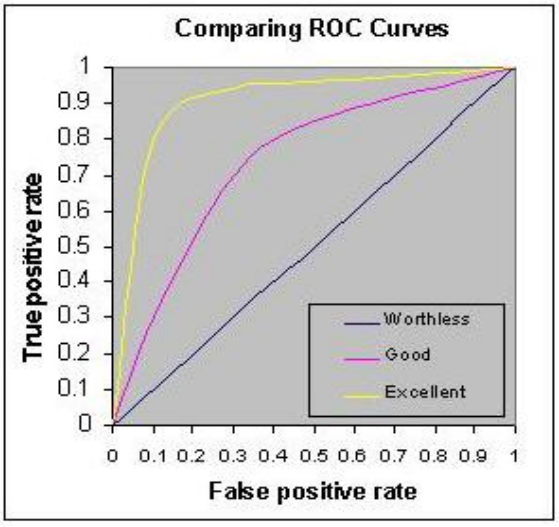
$$RE = (n_{12} + n_{21}) / N$$

Curba ROC

Clasificarea se face funcție de o valoare de prag, care poate fi un parametru complex determinat de algoritmul analizat. Indicatorii de estimare a calității au valori dependente de această valoare de prag: o valoare mai scăzută decât media va fi mai „îngăduitoare", acceptând mai lejer includerea unui subiect în clasa K+, caz în care se produce o creştere a sensibilității, reducându-se numărul de fals negativi. Din pacate, în acelaşi timp o parte din n_{22} trec în n_{21}, crescând numărul de fals pozitivi şi scăzând pe această cale specificitatea. Această relație între sensibilitate şi specificitate impune o „alegere optimă" a pragului, funcție de criteriile care sunt importante în anumite situații concrete şi modul cum se calculează *funcția de cost* a unei clasificări greşite. În funcție de scopul urmărit în studiu, vom căuta o sensibilitate mai ridicată (chiar dacă ştim că va creşte numărul de fals pozitivi), iar în altele vom urmări o specificitate crescută.

Concluzii interesante se pot trage dacă se urmăreşte grafic relația între sensibilitate şi specificitate. Se reprezintă de obicei

SN = f (1 – SP)



Curba ROC

Graficul obținut se numeşte curba ROC (*Receiver Operator Characteristic*). Acest grafic are câteva proprietăți interesante:

- dacă folosim un criteriu de clasificare fără putere de discriminare (de ex. un scor calculat absolut arbitrar), atunci curba ROC ar coincide cu prima bisectoare a planului, aria de sub curbă reprezentând 50% din total
- pe de altă parte, dacă am avea un criteriu perfect (nici un fals pozitiv sau negativ, indiferent de pragul P), atunci aria de sub curba ROC va fi 100%
- uzual curba ROC arată ca în fig.1.4.; aria de sub curba ROC este un indicator global satisfăcător pentru calitatea clasificatorului folosit şi se numeşte coeficientul c [46, 47, 48, 49, 50], http://gim.unmc.edu/dxtests/.

Coeficientul c

Aria de sub curba ROC se mai numeşte şi „coeficientul c" şi reprezinta cel mai important indicator al acurateței predicției (scorului).

Se acceptă următoarea scară în funcție de coeficientul c [http://gim.unmc.edu/dxtests]:

0,91 – 1,00 = excelent

0,81 – 0,90 = foarte bine

0,71 – 0,80 = bine

0,61 – 0,70 = satisfăcator

sub 0,60 = slab.

Exemplu: Reluăm exemplul cu cei 4000 subiecți dintre care 100 au avut viroză. Presupunem că programul nostru de calculator a diagnosticat corect 90 dintre ei, însă a atribuit acelaşi diagnostic (viroză) şi la alți 50 de subiecți. Caracterizați programul de diagnostic.

Datele din text sunt prezentate sintetic în tabelul IV.5.

Tabel IV.5. Exemplu pentru calculul parametrilor unui clasificator

	prog. calc. K+	K-	
B real	90	10	100
B-	50	3850	3900
	-	-	4000

– fals negativi F - = 10
– fals pozitivi F + = 50
– sensibilitatea SN = 90 / 100 = 90%
– specialitatea SP = 3850 / 3900 = 97,4%
– acuratețea AC = 3940 / 4000 = 98,5%
– rata erorii RE = 60 / 4000 = 1,5%

8. ALEGEREA INVESTIGAȚIILOR

Calculatorul poate asista medicul în luarea deciziilor nu numai pentru stabilirea diagnosticului ci şi în alte acțiuni, una dintre acestea fiind alegerea investigațiilor, acțiune desigur corelată cu stabilirea diagnosticului. Deseori medicul este solicitat de pacient sau familia acestuia să recomande investigații nu întotdeauna necesare sau relevante, uneori scumpe şi inavazive. Principalele elemente care intervin în luarea deciziei de recomandare sau nu a unei investigații sunt:

– relevanța rezultatului pentru conduita terapeutică sau precizarea diagnosticului (se efectuează o serie de calcule probabilistice pe baza regulii lui Bayes iar apoi, în funcție de sensibilitatea şi specificitatea testului se estimează probabilitatea evoluției în cele două variante - cu sau fără rezultatul investigației)
– tipul de investigație - invazivă, neinvazivă

– costul investigaţiei
– efecte secundare, contraindicaţii, accidente.

9. OPTIMIZAREA TRATAMENTULUI

Chiar dacă aplicaţiile actuale acoperă în mică măsură aspectul optimizării tratamentului medical, specialiştii apreciază că în următoarele două decenii aceasta va deveni aplicaţia majoră a informaticii medicale. Terapia actuală se bazează pe existenţa unor scheme de tratament în care - în majoritatea cazurilor - pacienţii sunt încadraţi adoptându-se cam aceleaşi doze şi intervale de timp, cu variaţii mai curând calitative, în cazuri de complicaţii etc. Este un deziderat major al medicinei actuale terapeutice trecerea spre individualizarea tratamentului. Primul pas major se realizează prin realizarea unui "model individualizat al pacientului" pe calculator (cuprinzând caracteristicile sale relevante pentru aspectul analizat), asupra căruia se simulează diferite "variante de tratament", alegându-se varianta optimă.

Astfel de programe au şi fost realizate pentru optimizarea tratamentului tumorilor prin iradiere. Problema importantă care trebuie rezolvată în aceste cazuri este atingerea dozelor terapeutice în regiunea tumorală fără însă a afecta regiunile străbătute de radiaţii până în zona tumorală. Iradierea sub mai multe incidenţe permite ca în zona tumorală, prin efect aditiv să se cumuleze doza terapeutică, fără a depăşi limitele admise pentru celelalte regiuni. Programele de calculator folosite precizează incidenţele şi dozele pentru fiecare incidenţă.

10. DECIZII LA NIVEL DE ORGANIZARE SANITARĂ

O largă paletă de aplicaţii este deschisă pentru asistarea deciziei la nivele centrale privind:

- distribuirea resurselor în funcţie de priorităţi şi necesităţi
- estimarea necesarului de medicamente, echipamente, infrastructură, personal
- reacţii operative în caz de epidemii, calamităţi, accidente
- elaborarea politicii sanitare în planuri de scurtă şi lungă durată.

BIBLIOGRAFIE ŞI REFERINŢE

JH van Bemmel, MA Musen (eds). *Handbook of Medical Informatics*. Springer Verlag, Heidelberg, 1997

JH van Bemmel, F Gremy, J Zvarova (eds): *Medical decision marking: diagnostic strategies and expert system*. North Holland, Amsterdam, 1995

G.I. Mihalaş. *Diagnosticul asistat de calculator* (în: Progrese în medicină, editor: Gh. Gluhovshi), Helicon, Timisoara, 1997

G.I. Mihalaş. *Strategii de diagnostic asistat de calculator în medicina internă* (în: Interdisciplinaritatea medicinii interne, editor: I. Romoşan), Helicon, Timisoara, 1993

Partea a V-a

SISTEME INFORMATICE MEDICALE

1. INFORMAȚIA MEDICALĂ

În capitolele anterioare am trecut în revistă numeroase aplicații ale calculatoarelor în domeniul medical (crearea bazelor de date, prelucrări statistice, achiziția şi prelucrarea biosemnalelor şi imaginilor medicale, diagnostic asistat etc.), privite însă ca aplicații punctuale, fără să acordăm atenție deosebită modului în care acestea se integrează în ansamblul activităților medicale. În acest capitol vom arbora o privire sintetică pentru a aborda activitățile din domeniul medical ca pe un sistem, cu numeroase elemente structurale, între care circulă informații şi vom urmări măsura în care tehnica de calcul poate sprijini aceste activități precum şi particularitățile acestui sistem în ansamblul său.

Şi deoarece elementul central urmărit aici este "informația medicală", vom începe printr-o privire sintetică asupra acestei noțiuni, pentru a putea defini conceptul de "sistem informatic".

Prima tentație în a defini noțiunea de "informație medicală" ar fi de a-i limita sfera de cuprindere la informații care privesc aspectele medicale din activitatea de ocrotire a sănătății. Vom extinde însă sfera acestei noțiuni astfel încât să cuprindă orice informație care apare în cursul activităților medicale, atât cele directe cât şi conexe. Să trecem în revistă tipurile de activități, încât să putem estima tipurile de informații care apar în diverse locuri şi momente, pentru a putea analiza apoi cum această informație circulă în sistemul medical / sanitar.

1.1. TIPURI DE ACTIVITĂȚI

a) Activități medicale directe - reprezentate tipic de "consultația medicală". Pentru a analiza tipurile de informații vehiculate, vom distinge câteva faze / acțiuni:

i^0 - stabilirea diagnosticului - faza în care medicul foloseşte două categorii de informații

- **date:** un ansamblu de informații cu caracter individual, cuprinzând: elementele culese în anamneză, datele de laborator, semnale, imagini
- **cunoştințe:** anasamblul de informații generale pe care le achiziționează medicul în cursul pregătirii sale profesionale (prin instruire, experiență clinică, documentare, cercetare)

ii^0 - tratament - fază în care medicul urmăreşte rezultatele terapiei propuse stabilindu-se un permanent schimb de informații între medic şi pacient.

iii^0 - nursing - denumire sub care acoperim toate activitățile privind îngrijirea pacienților.

b) Asigurarea logistică a activității medicale

Cadrul în care se desfăşoară orice activitate necesită o activitate organizatorică, administrativă şi managerială, mai simplă sau mai complexă, în funcție de specificul activității şi de dimensiunea sistemului. Vor fi desigur deosebiri între activitățile manageriale la nivel de circumscripție sanitară sau la nivel de spital. Aceste activități "conexe" cuprind cel puțin două elemente principale:

- activități de administrare a unității
- activități financiar contabile.

Trebuie menţionat că uneori aceste activităţi ocupă un procent însemnat din timpul consumat de medic în activitatea sa de ansamblu.

c) Integrarea în contextul social. Activitatea medicală nu este o acţiune izolată, cu scop în sine, ci face parte din ansamblul activităţilor dintr-o societate, astfel încât rezultatele activităţii medicale trebuie să fie vizibile la nivelul societăţii. Acest lucru se atinge pe mai multe căi; vom accentua aici însă acţiunea de centralizare a datelor medicale, prin care se raportează ierarhic datele sintetice cu ajutorul cărora se obţine o imagine de ansamblu asupra activităţilor din domeniul ocrotirii sănătăţii la nivelul unei comunităţi / societăţi.

d) Educaţia medicală - este o acţiune de importanţă deosebită, fiind veriga esenţială în transmiterea informaţiei medicale condensate sub forma de "cunoştinţe" medicale. Ea cuprinde:

i^0 - învăţământul cadrelor medicale:

- medici
- cadre medicale
- învăţământ postuniversiatr, educaţie continuă

ii^0- educaţia pacienţilor - element deosebit de important pentru anumite categorii de pacienţi (diabetici, gravide, astmatici, etc); acestui aspect i se acordă actualmente o atenţie specială şi se realizează numeroase programe pentru categorii extinse de pacienţi.

e) Documentarea medicală

În mod obişnuit, după încheierea studiilor sursa principală de informare devine documentarea din cărţi şi reviste de specialitate. Acestora li se adaugă acum metodele computerizate, folosind fie revistele publicate pe compact-disc (CD), exemplul tipic fiind sistemul MEDLINE elaborat de National Library of Medicine din Bethesda (SUA), fie conectarea pe Internet, cu acces la diferite biblioteci de specialitate din lume, unele dintre acestea asigurând acces gratuit. Toate aceste forme de completare şi actualizare a cunoştinţelor, împreună cu diverse cursuri post-universitare, se încadrează în conceptul de *educaţie medicală continuă*.

f) Cercetarea medicală

Cantitatea totală de informaţie este în creştere rapidă şi acest lucru se datorează dezvoltării deosebite a cercetării. Tehnologia informaţională contribuie din plin la această creştere rapidă inclusiv a cercetării medicale. Actualmente în toate instituţiile de învăţământ superior medical se desfăşoară şi o intensă activitate de cercetare, integrată de fapt în ansamblul activităţilor medicale.

1.2. STRUCTURA SCHEMATICĂ A FLUXULUI INFORMAŢIONAL

a) Schema fluxului informaţional

Vom prezenta în continuare tipurile de informaţii şi conexiunile de transfer a informaţiilor legate de activitatea medicală (fig. V.1).

Poziţia centrală în schemă o ocupă axa PACIENT - MEDIC care reprezintă activitatea medicală primară şi generează toate celelalte acţiuni. Medicul culege de la pacient informaţia medicală sub formă de date, care cuprind atât elemente descriptive din anamneză cât şi alte date: rezultate de laborator, semnale, imagini etc. Aceste date au caracter individual. Ele sunt interpretate de medic pe baza cunoştinţelor sale de specialitate, obţinute prin educaţie, documentare (cărţi, reviste, mijloace informatizate), experienţă clinică, eventual şi cercetare. Interpretarea datelor conduce la stabilirea unui

diagnostic şi elaborarea unui plan terapeutic care se aplică pacientului. Efectele tratamentului sunt urmărite de către medic închizându-se astfel un prim ciclu în care "noua" stare a pacientului privită ca "feed-back" în circuitul informaţional va determina o "nouă" decizie a medicului (aici atributul "nouă" reprezintă un nou moment, nu neapărat o **altă** stare sau decizie). Conform teoriei sistemelor, un sistem cibernetic în care există un ciclu cu legătura inversă este considerat un sistem reglabil (controlabil).

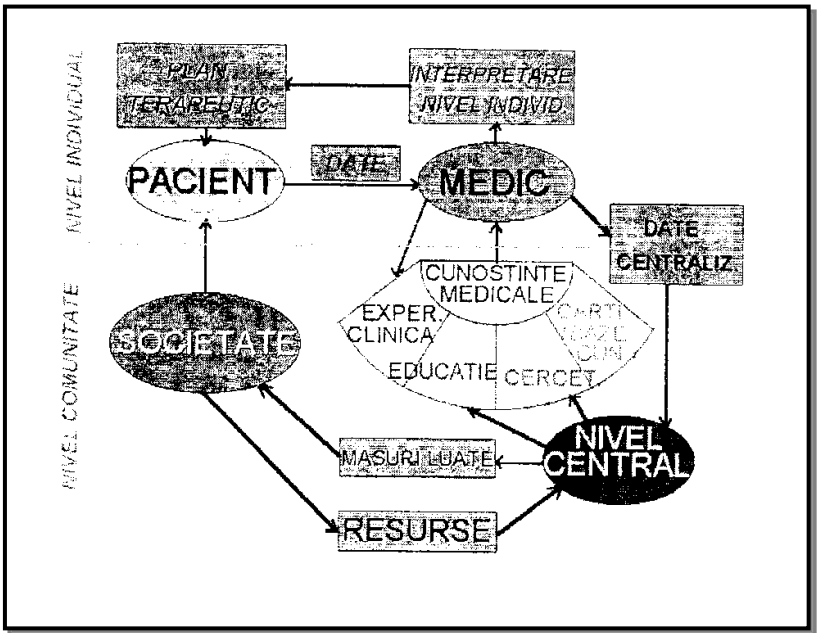


Fig. V.1. Schema fluxului informaţional în activitatea medical

În ciclul descris aici datele au un "nivel individual", acest lucru fiind marcat în schemă printr-o linie întreruptă care desparte porţiunea superioară de cea inferioară a figurii.

Caracterul individual al informaţiilor se pierde atunci când datele se centralizează pentru mai mulţi pacienţi, de la mai mulţi medici, fiind supuse unor prelucrări statistice la NIVEL CENTRAL. Se obţine o imagine de ansamblu asupra stării de sănătate a populaţiei, aceasta fiind o informaţie medicală la "nivel de comunitate". Cunoaşterea situaţiei la nivelul societăţii permite luarea unor măsuri de îmbunătăţire a activităţii de ocrotire a sănătăţii prin: măsuri profilactice, vaccinări, eventual acţiuni deosebite (în epidemii), orientarea unor fonduri spre anume activităţi achiziţionări de medicamente, echipamente etc. Măsurile luate se răsfrâng asupra întregii SOCIETĂŢI, implicit asupra pacienţilor ca indivizi. Acesta poate fi considerat ca un al doilea ciclu al fluxului informaţiei medicale.

Circuitul se închide şi prin influenţa pe care o poate avea informaţia la nivel central cu efect asupra creşterii/îmbunătăţirii bazei de cunoştinţe, cum ar fi orientări ale unor teme de cercetare, dezvoltarea învăţământului etc. Acesta ar fi al treilea ciclu al fluxului informaţional. Să nu trecem cu vederea că toate măsurile la nivel central depind de resursele disponibile pentru astfel de acţiuni, care - la rândul lor - sunt distribuite de către societate.

Se observă din schemă caracterul complex al transferului de informaţie în activitatea medicală, numeroasele legături şi tipurile diferite de informaţii.

b. Nivele de organizare a activităţii medicale

Putem distinge în activitatea medicală patru nivele de organizare:

- asistenţa medicală primară - asigurată la nivelul circumscripţiilor sanitare de către medici de medicină generală şi / sau medici de familie
- asistenţa medicală de specialitate - asigurată la nivelul cabinetelor şi / sau clinicilor de specialitate, având şi suportul unor servicii paraclinice
- spitalul ca unitate de organizare a asistenţei medicale de specialitate
- nivele centrale-decizionale: direcţii sanitare, ministere, legături cu organisme internaţionale (Organizaţia Mondială a Sănătăţii).

1.3. SISTEM INFORMAŢIONAL, SISTEM INFORMATIC

După ce am urmărit schema fluxului informaţional în activităţile medicale, putem da acum două definiţii:

a. Sistemul informaţional reprezintă un ansamblu de unităţi structurale între care are loc un schimb de informaţii.

b. Sistemul informatic reprezintă partea din sistemul informaţional care cuprinde utilizarea calculatoarelor.

Procentul ocupat de sistemele informatice medicale în ansamblul sistemelor informaţionale variază de la o ţară la altă şi de la un nivel la altul. Peste tot însă se constată o marcată tendinţă de creştere, actuala generaţie de studenţi medicinişti din România fiind cu siguranţă printre cei care vor contribui la creşterea acestui procent în ţara noastră.

2. SISTEME INFORMATICE ÎN ASISTENŢA MEDICALĂ PRIMARĂ

În cele ce urmează vom particulariza aspectele privind fluxul informaţional în cadrul asistenţei medicale primare.

Asistenţa medicală primară constituie primul contact al unui pacient cu sistemul medical. În ţara noastră asistenţa primară este asigurată în cabinete private de medicină generală sau medici de familie.

2.1. ACTIVITĂŢI LA NIVELUL UNITĂŢILOR DE ASISTENŢĂ MEDICALĂ PRIMARĂ

a) activităţi medicale propriu-zise, cuprinzând: consultaţii, vizite, urmărirea unor categorii speciale de pacienţi (bolnavi cronici, gravide, copii sub un an), urgenţe

b) documentare

c) organizare şi management la nivelul circumscripţiei (cabinetului).

Calculatoarele pot veni în facilitarea acestor activităţi. Vom prezenta în continuare tipurile de programe pe care ar trebui să le conţină un calculator al unui medic de medicină generală. Sunt trecute şi elemente (programe) privind aspecte care însă nu sunt definitivate în ţara noastră dar vor fi cu siguranţă introduse în viitor - ne referim aici în special la introducerea asigurărilor de sănătate, direcţie în care s-a acumulat multă experienţă în alte ţări şi în care aportul calculatoarelor este deosebit de important.

2.2. MODULELE SISTEMELOR INFORMATICE ALE ASISTENȚEI MEDICALE PRIMARE

1^0. Modulul de bază

Modulul de bază, obligatoriu în unele țări, cuprinde programele cel mai des folosite împreună cu fișierele aferente. Putem clasifica aceste componente în:

a) Fișiere de date medicale cuprinzând:

- registrele tuturor pacienților - de fapt a tuturor persoanelor în evidența circumscripției
- date "demografice", inclusiv date despre asigurările medicale ale pacienților
- posibilitatea grupării datelor pe familii (gospodării)

b) Programe de administrare a activității medicale

- registre (fişiere) pentru: consultații, vizite, teste de laborator
- elaborarea unor documente financiare simple - note de plată
- programe simple pentru diferite centralizări (medicamente etc.)

c) Programe utilitare ce asigură desfăşurarea unor operațiuni ca:

- asigurarea protecției datelor
- lucrul în partiție (când sunt mai multe terminale: în cabinetul medicului, la asistenta de recepție, în sala de laborator/investigații, în sala de tratamente etc.)
- salvarea datelor ("back-up") - o operațiune deosebit de importantă: este obligatoriu ca zilnic să se copieze toate fişierele de date pentru a putea fi restaurate în cazul unor defecțiuni ale sistemului de calcul.

2^0. Modulul medical

Are în poziție centrală fişa de observație, care este "placa turnantă" cu rol cheie în aplicațiile medicale (acest modul se mai întâlneşte sub numele **EPR** - "electronic patient record", sau CPR - "computer - based patient record").

a) **Fişa de observație** conține:

- date de identificare (cod personal, nume-prenume-adresa) şi alte date personale
- date medicale - care se trec grupate cronologic şi cuprind:
  - = antecedente personale şi heredo - colaterale (istoric)
  - = date ale examinării (ex: puls, presiune arterială, alte observații)
  - = rezultate de laborator (hemogramă, examen de urină, etc.)
  - = semnale (ECG, EEG etc.)
  - = imagini (radiografie, scintigrafie etc.)
  - = prescripții de medicamente, alte tratamente
  - = trimiteri spre asistența de specialitate.

Din punct de vedere al formei de înscriere, datele pot fi grupate în:

- date sub formă de text
- date numerice (cu precizie prestabilită)
- date codificate
- semnale şi imagini (în aceste situații se foloseşte termenul de format "multimedia")

b) **Codificarea** - este o operațiune frecvent întâlnită pentru o prezentare sub formă prescurtată a unor informații. Ea a fost introdusă inițial pentru evidența cauzelor de deces şi extinsă ulterior pentru descrieri complexe. Ansamblul de coduri folosite pentru un scop anume formează un nomenclator. Cele mai răspândite sisteme de codificare sunt:

i^{o} - ICD10 (International Classification of Diseases) versiunea 10-a, cu cea mai largă răspândire - este un sistem uniaxial, folsind un cod după un singur criteriu - diagnosticul

ii^{oo} - SNOMED (**S**ystematized **No**menclature of Human and Veterinary **Med**icine); este un sistem multiaxial, cu coduri separate pentru topologie, etiologie, morfologie, boală etc.

c) **Prescripția medicamentelor**

Elaborarea prescriptiei de tratament, este asistată de programele mai recente într-o formă destul de dezvoltată, având posibilitatea de a avertiza medicul de anumite situații, de genul:

- un medicament cu acțiune similară mai este inclus în tratament
- medicamentul respectiv este contraindicat în... (de ex. sarcină, etc.)
- medicamentul respectiv nu este plătit de firma de asigurări, etc.

d) **Trimiteri**

Modulul medical conține şi programe pentru editarea şi tipărirea scrisorilor de trimitere pentru diferite analize către cabinetul de specialitate sau spitale.

3^{0}. Modulul programări

Cu consecințe importante în economisirea timpului atât al pacienților cât şi al medicului; conține:

- evidența programărilor de consultații
- planificarea vizitelor
- programarea unor acțiuni speciale (vaccinări etc.)
- redactarea unor scrisori privind programările

4^{0}. Modulul farmacie

Unitățile de asistență primară au de obicei şi o dotare cu o serie de medicamente strict necesare şi pentru urgențe; ca o uzanță de activitate, cu ocazia consultației (vizitei) se pot administra deja unele medicamente, urmând ca restul să fie achiziționate conform rețetei; acest lucru este foarte important în circumscripțiile rurale. Modulul farmacie conține de obicei:

- evidența medicamentelor în stoc
- elaborarea comenzilor
- liste de gratuități / compensații
- lista furnizorilor
- date cumulative privind medicamentele prescrise într-un anumit interval de timp.

Observație: în multe țări casele de asigurări stabilesc "plafoane" de cheltuieli pentru tratamente, care să nu fie depăşite de către medici, de aceea, în mod uzual un doctor doreşte să-şi cunoască în orice moment nivelul la care a ajuns, comparativ cu plafonul prevăzut.

5^{0}. Modulul financiar

În afară de un program financiar-contabil inclus în modulul de bază, medicul generalist necesită un modul dedicat care să-i satisfacă integral necesitățile administrării financiare a unității, cuprinzând programe pentru:

- note de plată privind activitățile prestate
- statele de plată ale personalului

- registrul contabil, inclusiv impozite
- corespondența financiară.

6^0. Modulul de comunicație

Unitățile medicale nu sunt entități izolate, având poziții bine precizate într-o întreagă rețea de asistență medicală; totodată are legături şi cu unități aparținând altor rețele. Se apreciază că nivelul actual de comunicare este nesatisfăcător, datele de interes pentru un medic (cum ar fi cele privind tratamentele aplicate unui pacient într-o unitate specializată) sosind deseori incomplete şi cu întârziere. Este evident ca o conexiune a unităților de asistență primară cu cele specializate (spitale) ar îmbunătăți cu mult situația, oferind posibilitatea transferului rapid şi complet de informații.

7^0. Dezvoltări ulterioare

Deşi satisfac o bună parte din necesități, programele actuale pot fi încă îmbunătățite prin adăugarea unor facilități suplimentare:

- documentarea asistată – tehnici de data mining aplicate pentru resurse *web*
- adăugarea unor module de sisteme expert pentru asistarea deciziei medicale, în special diagnosticul (în momentul de față sistemele expert au o aplicare încă destul de restrânsă şi sunt cel mai adesea utilizate în clinici de specialitate).
- ridicarea nivelului de standardizare a fişei de observație.

3. SISTEME INFORMATICE CLINICE

Caracteristica esențială a sistemelor informatice clinice este că sunt orientate pe pacient, adică fişa pacientului este **documentul primar** şi toate relațiile şi conexiunile între departamente se fac cu referire la pacient. De aceea identificatorul pacientului este un element important pentru regăsirea facilă şi urmărirea datelor.

3.1. STRUCTURA ASISTENȚEI SPECIALIZATE ÎN CLINICI

Un pacient trimis de la nivelul asistenței medicale primare la nivelul asistenței de specialitate va intra în evidența unui departament clinic, însă pentru obținerea unei imagini complete asupra stării sale se apelează la o serie de servicii disponibile în departamente paraclinice. În principiu putem considera că un sistem policlinic conține două categorii de departamente: clinice şi paraclinice. Pentru un pacient se poate apela la serviciile oricărui departament paraclinic, iar aceste departamente paraclinice servesc toate departamentele clinice. Putem enumera principalele departamente de specialitate:

a) Departamente clinice

- medicină internă, subdivizate în: cardiologie, nefrologie etc.
- chirurgie, subdivizate la rândul lor
- pediatrie
- monitorizări
- psihiatrie
- neurologie
- boli infecțioase etc.

b) Departamente paraclinice şi servicii

- radiologie şi imagistică
- medicină nucleară
- explorări funcționale

- laborator clinic
- endoscopie
- laborator anatomo-patologie
- terapie
- farmacie etc.

3.2. OBIECTIVE GENERALE ALE SISTEMELOR INFORMATICE CLINICE

a. Planificarea îngrijirii şi intervenţiilor asupra pacienţilor.

b. Gestiunea datelor pacienţilor:

- achiziţia, stocarea şi regăsirea datelor (referitoare la anamneză, date de laborator, biosemnale, imagini etc)
- verificarea şi codificarea datelor
- prelucrarea datelor
- prezentarea integrată - în această direcţie există o avalanşă de programe ce propun numeroase variante ce pot oferi sintetic datele esenţiale şi cu acces uşor şi rapid la orice alte elemente, cuprinzând inclusiv imagini şi grafice.

c. Asistarea deciziei medicale - prin programe de:

- diagnostic asistat
- optimizare a terapiei
- simulări de evoluţii, inclusiv simulări de intervenţii pe modele.

d. Obiective educaţionale - de exemplu sfaturi pentru pacienţi.

e. Monitorizări şi urmărire - estimarea evoluţiei stării pacienţilor este un obiectiv esenţial al asistenţei medicale şi permite reacţia oportună pentru modificarea tratamentului după necesităţi; de aceea programele de calculator trebuie să permită obţinerea rapidă şi facilă a datelor solicitate prezentate într-o formă uşor interpretabilă. Adăugăm aici ca o categorie specială programele folosite pentru monitorizarea în terapia intensivă.

f. Raportare - activitatea medicală presupune redactarea periodică a unor rapoarte cuprinzând date sintetice ale activităţii, din care se va estima la nivel central starea de sănătate a populaţiei în vederea adoptării celor mai potrivite măsuri pentru îmbunătăţirea activităţilor de ocrotire a sănătăţii şi asistenţă medicală. Redactarea acestor rapoarte, care este o acţiune consumatoare de timp este mult uşurată prin utilizarea unor programe pentru:

- centralizarea datelor
- analiza statistică
- generare de rapoarte (există chiar forme standard care pot fi elaborate periodic).

g. Evaluarea calităţii asistenţei medicale şi a rezultatelor obţinute - stocarea uşoară a unui număr mare de date, regăsirea lor rapidă şi prelucrarea comodă permite aprecierea ori de câte ori este nevoie - a evoluţiei bolilor (în special în bolile cronice sau congenitale).

Putem include aici şi o altă categorie de programe, care permit o estimare realistă a calităţii asistenţei medicale pentru îmbunătăţirea planificării activităţilor şi resurselor în viitor şi chiar pentru orientarea unor activităţi de cercetare.

3.3. OBIECTIVE SPECIFICE ALE SISTEMELOR INFORMATICE ÎN DEPARTAMENTE CLINICE

Fără a avea pretenţia la o prezentare exhaustivă a aplicaţiilor calculatoarelor în fiecare specialitate, deoarece am prezentat anterior o serie de obiective generale, ne vom limita la o enumerare succintă a unor aplicaţii specifice, întâlnite mai des în anumite clinici.

a) Medicină internă

- Cardiologie: una din disciplinele clinice cu exprimările cele mai exacte, cu modele matematice dezvoltate şi mărimi mai uşor de cuantificat, folosind investigaţii destul de precise. Sunt tipice prelucrările de semnale ECG şi echocardiografia, precum şi de imagini (angiografie coronariană şi scintigrafie cardiacă)
- Boli metabolice: este specifică urmărirea pe lungă durată; pentru diabetici s-au creat o serie de programe de educaţie a pacienţilor
- Hematologie: s-au creat registre de hemofilie, cu date detaliate privind testele şi simptomele şi care reamintesc pacienţilor programările la consultaţii
- Nefrologie: baze de date internaţionale cu liste de priorităţi pentru transplant renal precum şi programe de telecomunicaţie între staţiile de dializă la domiciliu şi spital pentru monitorizarea pacienţilor cu insuficienţă renală cronică
- Gastroenterologie: prelucrări de imagini endoscopice, înregistrări multimedia (cu secvenţe video), regăsirea imaginilor şi compararea lor.

b) Chirurgie -aplicaţii pentru:

- planificarea operaţiilor
- pregătirea şi controlul intervenţiilor chirurgicale
- monitorizarea pacienţilor în timpul operaţiilor
- simularea unor operaţii prin tehnici de "realitate virtuală" - programe deosebit de utile pentru pregătirea viitorilor specialişti.

c) Oncologie - s-au realizat programe speciale pentru:

- codificări specifice (ONCOTOP)
- asistarea proiectării terapiei cu radiaţii
- elaborarea protocoalelor de chimioterapie
- centralizarea specifică pentru "Registrul naţional de cancer"
- prelucrări statistice specifice, inclusiv analiza supravieţuirii şi compararea tratamentelor.

d) Obstetrică:

- urmărirea sarcinii
- prelucrarea ultrasonocardiogramelor fetale
- educarea pacientelor
- estimarea calităţii îngrijirii gravidelor
- monitorizarea în timpul travaliului.

e) Pediatrie

- baze de date pentru prematuri, cu programe speciale de urmărire a evoluţiei creşterii
- depistarea precoce şi urmărirea bolilor congenitale
- alte aplicaţii sunt similare cu cele enumerate la medicină internă sau chirurgie, însă aplicate specific pentru copii de diferite vârste.

f) Psihiatrie

- baze de date
- programe de interpretare a unor teste specifice
- sisteme de diagnostic asistat (variabilitate destul de largă a diagnosticului).

g) Neurologie - domeniu foarte exact în stabilirea diagnosticului - programe pentru:

- estimarea gradului de disabilitate (scoruri)
- fişă de observaţie specifică
- controlul terapiei.

h) Monitorizări - cu programe specifice pentru diverse tipuri; se manevrează foarte multe date; calculatoarele folosite au plăci de achiziţie de semnal, cu mai multe canale; exemple de tipuri de monitorizări asistate de calculator - în:

- unităţi coronariene
- terapie intensivă / anestezie - reanimare
- urmărirea pre / post - operatorie
- monitorizarea perinatală
- administrarea bazei de organe pentru transplant.

3.4. OBIECTIVE SPECIFICE ÎN DEPARTAMENTE PARACLINICE ŞI SERVICII

a) Explorări funcţionale - se înregistrează diverse semnale biologice:

- explorări respiratorii
- ECG, EEG, EMG
- investigaţii în efort, etc.

b) Radiologie şi imagistică - au specific obţinerea de imagini şi necesită calculatoare cu mare capacitate de stocare, memorie mare şi viteză ridicată de transfer a datelor:

- radiografie (radioscopie)
- CT (computer - tomografie)
- RMN (rezonanţă magnetică nucleară)
- PET ("positron emission tomography")
- imagini echografice.

c) Laboratorul clinic - o serie de particularităţi pot fi menţionate aici:

- identificarea probelor prin coduri cu bare ("bar codes" - care oferă o mare operativitate) sau dispozitiv de citire OCR (*Optical Character Readers*)
- automatizarea comenzii (solicitării) de analize - în cazul în care este un pacient internat aceasta poate fi transmisă prin reţea şi pe aceeaşi cale pot fi primite şi rezultatele - inclusiv cu precizarea metodei (unele teste pot fi realizate prin mai multe metode şi există uşoare diferenţe între domeniile "normale" acceptate în funcţie de metodă)
- multe aparate de laborator permit o conectare la calculator şi au grad înalt de automatizare al procesării.

d) Laboratorul de patologie - cu rol important în diagnoza a două tipuri de probe:

- pe probe bioptice de la pacienţi
- pentru diagnoza post mortem.

În activitatea medicului patologist un rol important îl joacă "experienţa" câştigată prin "citirea" unui număr imens de lame; deseori medicul patologist apelează la cărţi şi atlase cu diverse imagini şi le compară cu cazul real; acestea pot fi acum furnizate de calculator; de asemenea s-au realizat unele sisteme expert care apelează la o serie de parametri numerici (număr de mitoze, dimensiunea şi forma nucleilor, conţinutul de AND).

e) Farmacia

Sistemul informational al unei farmacii de spital, care asigură servicii pentru toate celelalte departamente are numeroase sarcini ce pot fi grupate în:

i^0 - activităţi legate de asistenţa medicală:
- evidenţa tuturor reţetelor servite
- verificarea prescripţiilor
- furnizarea de informaţii la zi doctorilor şi asistentelor
- prepararea reţetelor magistrale

ii^0 - activităţi logistice:
- aprovizionare
- evidenţă stocuri şi termene de valabilitate
- lista furnizorilor

iii^0 - activităţi managerial.

4. SISTEME INFORMATICE DE SPITAL (SIS)

Spitalul reprezintă unitatea tipică de organizare a asistenţei medicale de specialitate. În paragraful precedent am trecut în revistă sistemele informatice din departamentele implicate direct în asistenţa medicală: departamentele clinice, paraclinice şi servicii medicale. Spitalul constituie un sistem complex, care integrează atât activităţile medicale din departamentele clinice şi paraclinice, cât şi întregul lanţ de activităţi conexe (administrative, financiare şi manageriale).

4.1. TIPURI DE DATE ÎN SPITAL

Disponibilitatea informaţiilor, în special sub formă de "date" este un factor cheie în funcţionalitatea unui sistem atât de complex.

Trebuie să facem aici distincţie între două tipuri de date: orientate pe pacient, respectiv orientate pe spital.

a) Date orientate pe pacient - sunt datele primare ale departamentelor clinice şi paraclinice; din fişierele pacienţilor se construiesc celelalte fişiere - rapoarte, centralizări etc. Conţinutul datelor orientate pe pacient este dinamic şi în creştere prin dezvoltarea metodelor de investigaţie şi terapeutice.

b) Date orientate pe spital - care cuprind datele referitoare la întreaga activitate managerială şi financiar contabilă precum şi datele sintetice extrase din fişierele activităţii medicale directe.

4.2. CONCEPTUL DE SIS

Pentru a putea mai bine defini noţiunea de Sistem Informatic de Spital prin sfera sa de cuprindere să trecem în revistă principalele funcţiuni pe care trebuie să le îndeplinească:

- sprijinirea activităţilor zilnice la nivelul asistenţei medicale directe
- suport în planificarea acestor activităţi
- sprijin în acţiunea de control şi corecţie a activităţilor medicale
- extragerea informaţiilor cu caracter statistic-populaţional
- accesul la baza de date medicale pentru cercetarea clinică

a) Scopul SIS: utilizarea calculatoarelor pentru colectarea stocarea şi prelucrarea informaţiei privind asistenţa acordată pacienţilor, precum şi administrarea în toate activităţile legate de spital şi a satisfacerii cerinţelor funcţionale.

Crearea unui SIS asigură:

- o utilizare mai eficientă a resurselor (întotdeauna limitate!) disponibile pentru asistenţa acordată pacienţilor
- îmbunătăţirea calitativă a serviciilor oferite
- un sprijin operativ pentru nivelele centrale în vederea cunoaşterii stării de sănătate a populaţiei într-un teritoriu
- cadrul adecvat pentru învăţământul medical şi pentru cercetare.

b) Componentele SIS

- **Baza de date a pacienţilor.** Cum în centrul activităţilor medicale se găseşte **pacientul,** poziţia centrală în SIS o ocupă această bază de date. Ea trebuie creată astfel încât dezvoltarea ulterioară a tehnicii de calcul sau limbajelor să permită utilizarea ei în continuare.

- **Aplicaţiile -** reprezentate de diferitele programe de prelucrare a datelor din baza de date, pornind de la programele simple pentru introducerea şi modificarea datelor, prezentarea fişei pacientului, reprezentări grafice pentru evoluţia pacientului până la programe de centralizare pe zile, boli, tratamente, medicaţii, investigaţii precum şi analize statistice de diverse tipuri.

- **Sistemul de comunicaţie -**care cuprinde facilităţile de legătură între baza de date şi utilizatorii individuali; la acest nivel se poate realiza limitarea acccesului diferiţilor utilizatori.

- **Terminalele de lucru -** care se găsesc distribuite în clinici, laboratoare, servicii şi birouri.

Schematic putem reprezenta componentele SIS ca în figura V.2.

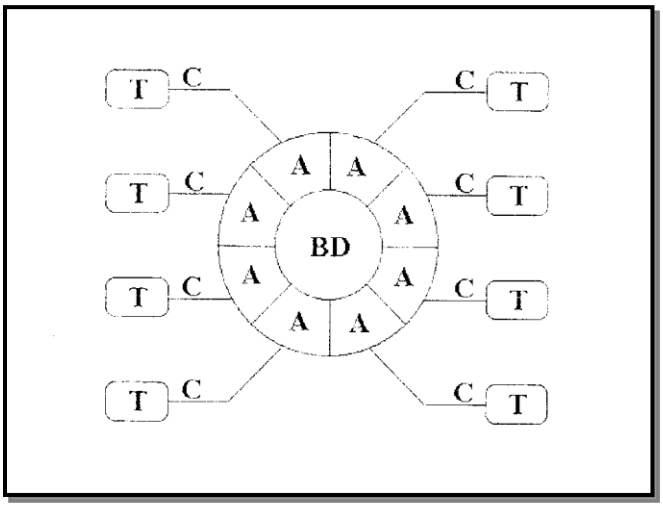


Figura V.2. Componentele sistemului informatic de spital: utilizatorii de la terminalele de lucru T au acces la baza de date BD prin sistemul de comunicaţie C, folosind diferite aplicaţii A

4.3. ARHITECTURA UNUI SIS

Evoluția sistemelor informatice de spital a relevat posibilitatea unor abordări diferite, cele mai uzuale arhitecturi fiind:

a) Sisteme monolitice - concepute şi construite într-o viziune unitară inițială; au avantajul unei bune compatibilități între componente, însă dezavantajul de a fi scumpe ca investiție unitară; s-au dovedit a fi mai puțin flexibile şi mai greu de conectat la sisteme externe diferite sau de adaptat la creşteri neprevăzute.

b) Sisteme evolutive - care au apărut ca necesitate a adaptării permanente a arhitecturii la necesități; putem aici distinge două situații:

- extinderea sistemelor monolitice (sisteme evolutive de tip I) prin adăugarea de noi componente

- conectarea unor sisteme izolate (sisteme evolutive de tip II) care sunt cele mai frecvente. Foarte multe SIS actuale au apărut prin integrarea la un moment dat a unor sisteme departamentale izolate. Deşi există numeroase dezavantaje (deseori fişierele bazelor de date nu au concepție unitară), se folosesc limbaje diferite şi calculatoare diferite, marele avantaj al cheltuielilor mai reduse şi dezvoltările flexibile au făcut ca acest sistem să fie cel mai frecvent.

c) Sisteme distribuite - destul de asemănătoare ca idee cu sistemele evolutive de tip II, prin adoptarea în arhitectură a unor componente diferite, care pot chiar rula pe platforme diferite (cu diferite sisteme de operare) şi să comunice cu baza de date pe baza unor protocoale standard de comunicație.

4.4. STRUCTURA UNUI SIS

În schema din figura V.3 este prezentată structura unui SIS.

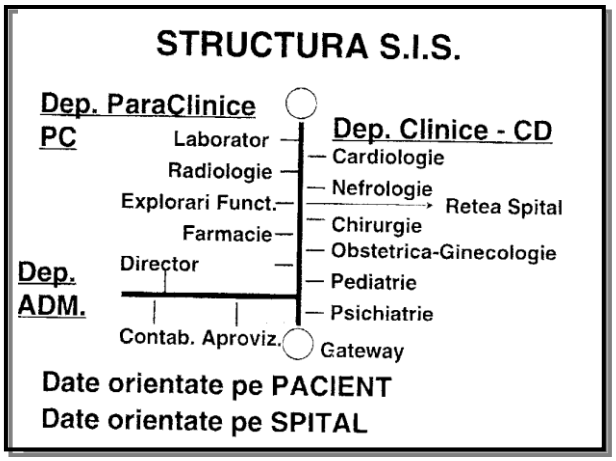


Figura V.3. Structura unui sistem informatic de spital

Un astfel de sistem conține **două magistrale de comunicație**:

a) - **magistrala de date medicale** - la care sunt conectate toate departamentele care lucrează cu date orientate pe pacient, atât departamentele clinice, notate DC (medicină internă: cardiologie, nefrologie, gastroenterologie, endocrinologie, boli metabolice, hematologie; departamente chirurgicale: chirurgie generală, urologie, neurochirurgie, chirurgie cardio-toracică, ORL, oftalmologie; departamente complexe: oncologie, obstetrică - ginecologie; pediatrie, psihiatrie, boli infecțioase (de obicei localizate în clădiri diferite - etc), cât şi departamentele paraclinice (laborator clinic,

radiologie şi imagistică, endoscopie, explorări funcţionale, laborator de patologie, medicină nucleară, anestezie etc.), departamente de monitorizare (terapie intensivă, unitate coronariană, reanimare, dializă), departamente de servicii medicale (farmacie, centrul de transfuzii, medicină legală, morga) etc.

b) - **magistrala de date de spital** - la care sunt conectate toate serviciile administrative, manageriale şi de suport logistic: blocul operator, blocul alimentar, serviciul tehnic/întreţinere, staţia de salvare, serviciul aprovizionare, serviciul personal, contabilitate-financiar şi conducerea spitalului.

4.5. INTEGRAREA SIS

Noţiunea de "integrare" în terminologia de aici reprezintă faptul că SIS nu este izolat (reţea locală), ci conectată la alte sisteme informatice, cu activităţi conexe, între SIS şi respectivele reţele.

În figura V.4 este prezentată o schemă cuprinzând conexiunile posibile ale unui sistem informatic de spital, cu alte reţele cu care există permanente schimburi de date. Schimburile între reţele sunt asigurate prin calculatoare de comunicaţie numite "Gateway", care permit cuplarea unor calculatoare cu sisteme de operare diferite (Unix, Windows, DOS, Apple).

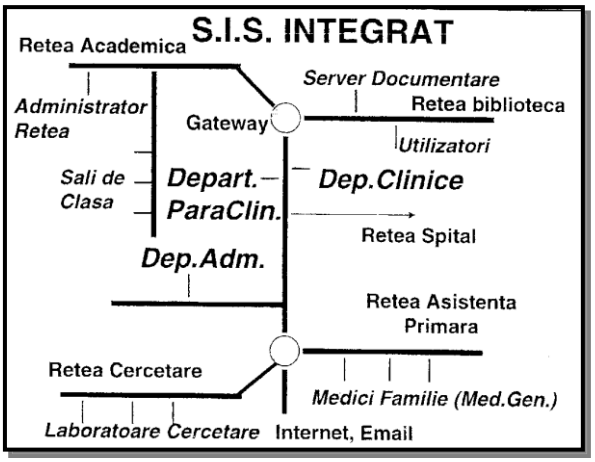


Figura V.4. Integrarea unui sistem informatic de spital

Sistemul integrat cuprinde următoarele componente:

a) SIS - Reţeaua de spital - care în contextul nostru constituie "coloana vertebrală" a întregrului sistem, şi este reprezentată prin porţiunea între cele două servere "gateway".

b) Reţeaua academică - marile unităţi spitaliceşti constituie şi o bază de învăţământ: atât a viitorilor medici şi cadrelor medii auxiliare (asistente, tehnicieni, etc.), cât şi învăţământului postuniversitar. Acesta are loc atât în forma tradiţională (rezidenţiat - formare a medicilor de specialitate), cât şi ca "educaţie continuă" pentru aducerea la zi a nivelului cunoştinţelor cadrelor medicale, în condiţiile actuale în care apar continuu noi cunoştinţe, metode, medicamente. Reţeaua academică permite utilizarea bazelor de date ale clinicilor şi programelor pentru procesul educaţional.

c) Rețeaua de bibliotecă - cu menirea de a asigura documentarea cadrelor medicale şi a studenților. Cel mai adesea o astfel de rețea are un "server" cu posibilități de a servi simultan multe sisteme de calcul conectate la Internet.

d) Rețeaua de cercetare - omniprezentă în spitalele universitare, dar şi în alte spitale - permite cuplarea diferitelor laboratoare de cercetare în rețeaua integrată; schimbul de informații este bidirecțional: laboratorul de cercetare necesită deseori date clinice în timp ce departamentele clinice doresc să aibă la dispoziție cât mai curând cele mai recente noutăți ştiințifice.

e) Rețeaua de asistență medicală primară. Trebuie să remarcăm mai întâi faptul că sistemele informatice ale asistenței medicale primare sunt deseori izolate, însă sunt deja create premizele ca astfel de rețele ce interconectează mai multe circumscripții / cabinete de medicină generală. Legarea acestor rețele la un sistem integrat oferă avantaje atât medicilor generalişti (prin accesul la datele pacienților cărora li se acordă asistența de specialitate, prin accesul la bazele de cunoştințe etc.) cât şi spitalului, pentru transfer rapid de date şi urmărirea ambulatorie lejeră.

4.5. EXEMPLE DE SIS

Până în prezent au fost realizate numeroase sisteme informatice de spitale. În literatura de specialitate sunt menționate câteva mai deosebite, în care s-au experimentat diferite soluții şi s-au creat de fapt standardele de construcție, astfel încat ele au putut fi actualizate pe paarcursul evolutiei tehnologice.

- Sistemul HISCOM - creat de firma Hiscom din Olanda - este un sistem complet, cu grad înalt de integrare, început în 1975
- Sistemul DIOGENE - creat la spitalul Cantonal din Geneva - cu arhitectură monolitică, construit în perioada 1971-1978
- Sistemul MGS (Massachusetts General Hospital) - creat la Boston, în perioada 1965-1970

Mai menționăm sistemele de la King′s College (Londra), Universitatea din Hanovra, Spitalul din Stokholm, Spitalul de Reabilitare din Texas, Spitalul Universitar din Tokyo. Sistemul informatic al rețelei de spitale VA (*Veterans' Administration*) din Statele Unite este aproape în întregime în formă electronică (96% *paperless*), fiind unul dintre exmplele de succes în domeniu.

O prima încercare în România a fost un proiect de realizare a unui SIS la Spitalul Fundeni din Bucureşti, demarat în anul 1980, dar nefinalizat. În prezent, un sistem informatic bine realizat, cu grad de complexitate deosebit este cel al Serviciului de Salvare din Bucureşti.

5. SISTEME INFORMATICE MEDICALE LA NIVEL CENTRAL

În fiecare țară există un specific al organizării ierarhice a activităților medicale. În principiu însă se respectă structura prezentată de noi, pornind de la asistența medicală primară asigurată în circumscripții / cabinete de medicină generală, spre cea de specialitate, având ca unitate fundamentală de organizare spitalul, conform schemei de flux informațional prezentată în figura III.4.1. Dinspre aceste unități care asigură asistența medicală directă, se centralizează date către aşa numitul nivel central. Aici se disting următoarele trepte specifice în România:

5.1. NIVEL TERITORIAL

DSJ (Directiile Sanitare Judetene) organizează şi supraveghează activităţile de asistenţă medicală la nivel teritorial - judeţ. La acest nivel se face centralizarea primară a datelor, fiind primul nivel de sinteză.

În plus, în judeţe există Case de Asigurări de Sănătate judeţene, ca filiale ale Casei Naţionale de Asigurări de Sănătate (CNAS) – ele raportează către CNAS, dar au şi un anumit grad de autonomie locală.

5.2. NIVEL NAŢIONAL

a) Ministerul Sănătăţii - centralizează la nivel naţional datele privind ocrotirea sănătăţii şi asistenţa medicală. În cadrul MS funcţionează Centrul Naţional de Statistică Sanitară care concentrează toate informaţiile şi redactează rapoartele de sinteză la nivel naţional. Pe acestea ministerul le analizează, le prezintă guvernului şi ia măsurile cuvenite pentru ridicarea calităţii ocrotirii sănătăţii şi asistenţei medicale.

De menţionat că reţeaua farmaceutică are o ierarhie teritorială paralelă reţelei de asistenţă medicală, datele fiind centralizate la Oficiul Central Farmaceutic.

b) Casa Nationala de Asigurari de Sanatate (CNAS) – primeşte informaţii şi finanţează serviciile de îngrijire a sănătăţii din: asistenţa primară la nivelul medicilor de familie, asistenţa în ambulatorii de specialitate, asistenţă în spitale (atât pentru afecţiunile acute, cât şi pentru cele cronice).

c) Alte ministere - în România, pe lângă reţeaua MS care satisface majoritatea acţiunilor de asistenţă medicală, există unele ministere care au reţele proprii de asistenţă medicală: Ministerul Transporturilor, Ministerul Apărării Naţionale şi Ministerul de Interne.

d) Organe Centrale. Există activităţi cu caracter medical (direct sau conex), desfăşurate şi în alte organisme:

- unităţi ale Ministerului Muncii şi Protecţiei Sociale (azile de bătrâni, inspectorate pentru handicapaţi etc.)
- unităţi ale Poliţiei Sanitar-Veterinare.

Unităţile de nivel central (judeţean sau naţional) au datoria de a asigura legătura cu diferite alte unităţi în probleme comune (de ex. alimentarea cu apă, colectarea şi depozitarea gunoaielor, diverse aspecte ecologice etc.).

Date DRG – baza informaţională pentru finanţarea spitalelor

Începand cu 2004, finanţarea spitalelor în România se face prin plata prospectivă bazată pe sistemul DRG (*Diagnosis Related Groups*). Sistemul a fost iniţial introdus experimental într-un singur spital (Cluj, 1999) printr-un proiect finanţat de USAID, apoi în 23 de spitale (2002) experienţa fiind apoi aplicată la nivel naţional.

Sistemul DRG reprezintă o schemă de clasificare a pacienţilor care permite relaţionarea tipurilor de pacienţi trataţi într-un spital (i.e. *case-mix*-ul) cu costurile cărora el trebuie să le facă faţă. Concepţia şi dezvoltarea sistemului a început la universitatea Yale la sfârşitul anilor 1960. Motivaţia iniţială pentru dezvoltarea lui a constituit-o crearea unui cadru care să permită monitorizarea calităţii şi a utilizării serviciilor în domeniul îngrijirii sănătăţii. El s-a extins treptat în SUA şi în 1983 a devenit sistemul de plată prospectivă la nivel naţional pentru toţi pacienţii *Medicare* (i.e. toţi cetăţenii de peste 65 de ani din Statele Unite).

În prezent, sistemul DRG se utilizează nu numai pentru pacienţii *Medicare*, ci este folosit ca metodă preferată de rambursare pentru majoritatea companiilor de asigurari şi, cu modificari, a fost introdus în Australia şi în mare parte din ţările europene. Evoluţia sistemului DRG şi utilizarea lui ca unitate de bază în plata spitalelor reprezintă o recunoaştere a rolului fundamental pe care *case-mix*-ul unui spital îl joacă în determinarea costurilor. Utilizarea altor caracteristici în stabilirea costurilor (statutul de spital universitar, numărul de paturi, etc.) a eşuat în găsirea unor explicaţii convingatoare atât privind diferenţele de costuri dintre spitale cât şi noţiunile de complexitate a cazurilor tratate. Sistemul DRG a fost primul sistem operaţional care a oferit mijloacele de a defini şi a cuantifica noţiunea de complexitate *case-mix*.

Termenul de complexitate *case-mix* este utilizat ca referinţă la un set de atribute ale pacientului care sunt inter-relaţionate dar distincte şi includ: severitatea bolii, prognosticul, dificultatea tratamentului, necesitatea intervenţiei şi intensitatea resurselor utilizate.

În sistemul DRG original există 25 de grupe diagnostice şi clasificarea pacienţilor în acestea se face pe baza diagnosticelor ICD (*International Classification of Diseases* recomandată de Organizaţia Mondială a Sănătăţii) prin care este descris cazul respectiv. Majoritatea ţărilor europene şi Australia au adoptat deja ICD-10 (versiunea 10 ICD) uneori modificată la nevoile proprii, în timp ce Statele Unite a rămas încă pe ICD-9, care are mai multe grupe diagnostice. În diferite ţări europene sistemul de clasificare şi finanţare a fost adaptat la cerinţele şi cultura instituţională proprii.

România a început cu sistemul american şi continuă cu cel australian, începand cu jumătatea anului 2007. O dificultate majoră o constituie dezvoltarea de valori relative locale pe baza datelor de costuri la nivel de pacient, precum şi organizarea şi dezvoltarea unei structuri de evaluare a calităţii serviciilor furnizate de spitale.

Un sistem de codificare mai complex este SNOMED, care permite o abordare ontologică a informaţiei medicale, dar încă nu este utilizat pe scară largă.

5.3. NIVEL INTERNAŢIONAL

Este tot mai evident că nici o ţară nu poate neglija contextul global în care este integrată şi multe probleme (inclusiv de ordin medical) sunt comune. Deschiderea graniţelor, circulaţia intensă, turismul etc. impun intensificarea comunicării internaţionale. În domeniul medical există un organism cu sediul la Geneva: OMS - Organizaţia Mondială a Sănătăţii, care are mai multe departamente şi care asigură comunicarea datelor şi informaţiilor medicale la nivel mondial. România este membră a OMS încă de la înfiinţare (1950).

Comunicarea la nivel internaţional este asigurată şi prin intermediul unor societăţi sau asociaţii internaţionale; de exemplu în domeniul informaticii medicale, România este membră atât a EFMI (Federaţia Europeană de Informatică Medicală) cât şi IMIA (Asociaţia Internaţională de Informatică Medicală).

Organismele internaţionale au rol informativ şi consultativ fiind deseori promotoarele unor proiecte preluate apoi la nivel naţional. În domeniul informaticii medicale problemele majore care se discută privind sistemele informatice medicale integrate sunt cele referitoare la protecţia datelor şi cele referitoare la standardizare, pe care le vom trece în revistă în cele ce urmează. Ca şi in celelalte probleme legate de îngrijirea sănătăţii, în informatica medicală Uniunea Europeană nu dă directive ci doar face recomandări statelor membre.

6. PROBLEME SPECIFICE ÎN SISTEME INFORMATICE

6.1. PROTECȚIA DATELOR

Realizarea oricărui sistem informatic (de fapt orice conexiune între calculatoare) ridică probleme legate de asigurarea confidențialitatii şi protecției datelor. Apare aici o contradicție: pe de o parte unul din scopurile pentru care se realizează sistemele informatice este chiar asigurarea **accesibilității** datelor, pe de altă parte datele medicale au un specific aparte, caracterul individual privat impunând ca aceste date să fie **confidențiale.** De asemenea, dorim să asigurăm **integritatea** lor, deci să nu fie afectat conținutul lor (fie accidental, fie intenționat). Termenii legați de aceste aspecte sunt următorii:

a) Confidențialitatea - datele medicale ale unui pacient sunt considerate confidențiale; accesul la ele trebuie deci să fie limitat la un număr redus de persoane.

Metode de asigurare a confidențialității: folosirea unor parole pentru a accesa fie întregul fişier (sau unele câmpuri) fie programele; codificarea identității pacientului.

b) Protecția datelor - reprezintă măsurile împotriva deteriorării accidentale - neatenție în manevrare, defecțiuni tehnice.

Metode uzuale de protecție: "salvarea" datelor şi programelor prin realizarea periodică (zilnică) a unor copii de siguranță (*back-up*) pe suport extern; de asemenea, se stabilesc nişte reguli stricte de operare şi evidență.

c) Securitatea datelor - reprezintă măsurile împotriva accesului sau deteriorării intenționate a datelor sau programelor.

Metode uzuale de securitate: introducerea unor parole de acces pentru diferite nivele sau chiar măsuri hardware - folosirea unor cartele, chei etc.

Asigurarea confidențialității şi integrității componentelor SIS impune stabilirea unor măsuri atât la nivelul conducerii spitalului cât şi la nivelul departamentelor. Ele pot fi grupate în trei categorii:

i^0 - măsuri *hardware* (sau de echipament):

- uneori echipamentele centrale sunt duplicate
- calculatorul central - într-o sală încuiată, cu acces limitat
- calculatoarele să permită accesul după identificarea persoanei cu o cartelă (magnetică, optică sau chip)
- instalație de avertizare a accesului neautorizat
- protecție împotriva inundațiilor
- instalații de aer condiționat

ii^0 - măsuri *software*:

- teste de verificare a programelor (cu date foarte variate)
- teste de validare a datelor introduse (atât la introducere cât şi cu anumite periodicități)
- identificarea utilizatorului - cu parole - şi a nivelelor de acces prestabilite: citire integral sau parțial; se recomandă modificarea periodică a parolelor
- păstrarea versiunii anterioare
- evidența actualizărilor în toate fişierele în care apar date ce se modifică

iii^0 - măsuri organizatorice care trebuiesc sistematizate într-un "Regulament al sistemului informatic" şi care trebuie să includă cel puțin:

- precizarea exactă (separarea) a sarcinilor

- prevenirea situațiilor în care prea multe privilegii aparțin unei singure persoane (situații care cresce riscul de abuz)
- clasificarea datelor în grupe diferite, cu acces depinzând de tipul de date
- legarea accesului la date de vechimea lor sau de originea lor
- elaborarea unor manuale de operare pentru fiecare funcție în sistem, cu precizarea procedurilor de urmat în diferite situații
- managementul autorizațiilor trebuie de asemenea elaborat ținând cont de structura organizatorică a unității.

6.2. STANDARDIZAREA

Termenii de standard şi standardizare, folosiți frecvent în tehnică păreau a fi greu adaptabili la activitatea medicală care manevrează numeroase noțiuni fuzzy (definite vag). Totuşi, în ultimul timp se discută din ce în ce mai intens despre standardizare pentru cel puțin două motive:

- asigurarea calității serviciilor prin instalarea unor norme precise pentru toate activitățile, asigurându-se o responsabilitate bine precizată a tuturor persoanelor implicate în ansamblul activităților
- posibilitatea schimbului eficient de informații între unități diferite - prin precizarea semnificației termenilor.

a) Definiții -după ISO (*International Standards Organization*).

Standardizare - operațiunea de stabilire a unor reguli de desfăşurare a unor acțiuni, privind probleme actuale sau potențiale, pentru atingerea unui grad optim de ordine într-un context dat.

Standard - este un document stabilit prin consens şi aprobat de un organism recunoscut, ce stipulează, pentru acțiuni comune şi repetate, reguli şi criterii pentru activități sau rezultatele lor, cu scopul de a atinge un grad optim de ordine într-un context dat.

Standardele din domeniul informaticii medicale permit interoperabilitatea între sistemele informaționale de sănătate.

b) Organisme naționale şi internaționale

- ONS - Oficiul Național de Standarde din România, Comitetul Tehnic 319
- CEN - Comitetul European pentru Standardizare (norme)
- CEN/TC251 - Comitetul Tehnic European pentru Informatică Medicală
- ANSI - Institutul Național American de Standarde
- ISO - Organizatia Internațională pentru Standarde.

c) Etapele lansării unui standard european

La nivelul CEN, în diferitele comitete tehnice se elaborează proiectele de standarde care trec prin următoarele etape:

ENV - pre-standard european: formă preliminară anunțată pentru a fi verificată pe o perioadă de 3 ani la nivele naționale

EN - standard european: un prestandard acceptat devine normă obligatorie; prevederile naționale care nu sunt în concordanță trebuie retrase pentru adaptarea deplină a standardului

CR - "CEN - raport": este prestandard la care nu s-a ajuns la consens, însă are prevederi atât de importante încât sunt făcute publice; totodată pot deveni CR alte documente normative care nu sunt propuse ca standarde

HD - "*Harmonization Document*" este un standard adaptat ca şi EN, dar permite unele variaţii naţionale pentru o perioadă de tranziţie.

d) Caracteristicile unui standard

Enunţarea unui standard trebuie să îndeplinească anumite condiţii, denumite precurtat "SMART":

. S - specific - obiectul să fie bine definit, clar, fără ambiguităţi

. M - măsurabil - acţiunile să poată fi măsurate şi exprimate cantitativ şi calitativ

. A - acceptabil de către instituţiile care îl utilizează

. R - realistic - să cuprindă acţiuni ce pot fi intreprinse practic

. T - "*time-related*" – acţiunile trebuiesc precizate în timp ca termene şi durate.

e) Standardul HL7

HL7 este o organizaţie internaţională, înfiinţată cu mai bine de 20 de ani în urmă, care dezvoltă standarde pentru schimbul de informaţii electronice în domeniul sănătăţii, precum şi de management şi integrare a acestor informaţii. HL7 nu dezvoltă *software*, ci specificaţii (e.g. *messaging standard*) care să permită interoperabilitatea – aplicaţii disparate să poată schimba cel puţin un set minimal de date clinice şi administrative.

Standardul **HL7 permite interoperabilitate**:

- tehnica – datele pot fi mutate din sistemul A în sistemul B
- semantica – asigură că sistemul A şi sistemul B înteleg datele în acelaşi fel
- de proces – permite ca activităţile organizaţiilor care găzduiesc sistemele A şi B să fie compatibile şi să se desfăşoare împreună.

BIBLIOGRAFIE ŞI REFERINŢE

RK Bali, AN Dwivedi (eds). *Healthcare knowledge management*. Springer, New York, 2007

JH van Bemmel, MA Musen (eds). *Handbook of Medical Informatics*. Springer, Heidelberg, 1997

A Bowling. *Research methods in health: investigating health and health services*. Open University Press, McGraw-Hill House, Maidenhead England, 2002

DRG Romania: *http://www.drg.ro/*

HL7 web site: *http://www.hl7.org*

ICD-10. International Statistical Classification of Diseases and Related Health Problems 10th RevisionVersion for 2007. *http://www.who.int/classifications/apps/icd/icd10online*

SNOMED. International Health Terminology Standards Development Organisation: *http://www.ihtsdo.org/*

L. Stoicu-Tivadar. *Sisteme informatice aplicate in sanatate*. Editura Politehnica, Timişoara, 2005